

UNCLASSIFIED



**Release 3.0**  
**September 2000**

# **INFORMATION ASSURANCE TECHNICAL FRAMEWORK**

Issued by: National Security Agency  
Information Assurance Solutions  
Technical Directors

**Disclaimer:**

This Information Assurance Technical Framework is the result of a collaborative effort by various organizations within the U.S. Government and industry. This document captures security needs and potential technology solutions for information systems and networks.

The information contained in this document is provided for information purposes only.

This is not a solicitation for procurement. Rather, this document is intended to facilitate the coordination of the information systems security needs of the U.S. Government and to offer security solution recommendations based on the collaborative efforts of the joint Industry/Government Information Assurance Technical Framework Forum.

UNCLASSIFIED

Report Documentation Page		
<b>Report Date</b> 01092000	<b>Report Type</b> N/A	<b>Dates Covered (from... to)</b> -
<b>Title and Subtitle</b> Information Assurance Technical Framework (IATF)	<b>Contract Number</b>	
	<b>Grant Number</b>	
	<b>Program Element Number</b>	
<b>Author(s)</b>	<b>Project Number</b>	
	<b>Task Number</b>	
	<b>Work Unit Number</b>	
<b>Performing Organization Name(s) and Address(es)</b> Booz Allen & Hamilton 8283 Greensboro Drive McLean, VA 22102	<b>Performing Organization Report Number</b>	
<b>Sponsoring/Monitoring Agency Name(s) and Address(es)</b> National Security Agency Information Assurance Solutions Technical Directors	<b>Sponsor/Monitor's Acronym(s)</b>	
	<b>Sponsor/Monitor's Report Number(s)</b>	
<b>Distribution/Availability Statement</b> Approved for public release, distribution unlimited		
<b>Supplementary Notes</b>		
<b>Abstract</b>		
<b>Subject Terms</b> IATAC COLLECTION		
<b>Report Classification</b> unclassified	<b>Classification of this page</b> unclassified	
<b>Classification of Abstract</b> unclassified	<b>Limitation of Abstract</b> UU	
<b>Number of Pages</b> 773		

<b>REPORT DOCUMENTATION PAGE</b>			<i>Form Approved</i> <b>OMB No. 074-0188</b>	
<small>Public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing this collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden to Washington Headquarters Services, Directorate for Information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302, and to the Office of Management and Budget, Paperwork Reduction Project (0704-0188), Washington, DC 20503</small>				
<b>1. AGENCY USE ONLY (Leave blank)</b>		<b>2. REPORT DATE</b> 9/1/2000	<b>3. REPORT TYPE AND DATES COVERED</b> Publication 9/1/2000	
<b>4. TITLE AND SUBTITLE</b> Information Assurance Technical Framework (IATF)			<b>5. FUNDING NUMBERS</b>	
<b>6. AUTHOR(S)</b> National Security Agency Information Assurance Solutions Technical Directors				
<b>7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)</b>  Booz Allen & Hamilton 8283 Greensboro Drive McLean, VA 22102			<b>8. PERFORMING ORGANIZATION REPORT NUMBER</b>	
<b>9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)</b>  National Security Agency Information Assurance Solutions Technical Directors			<b>10. SPONSORING / MONITORING AGENCY REPORT NUMBER</b>	
<b>11. SUPPLEMENTARY NOTES</b>				
<b>12a. DISTRIBUTION / AVAILABILITY STATEMENT</b> Approved for public release; distribution is unlimited			<b>12b. DISTRIBUTION CODE</b>  A	
<b>13. ABSTRACT (Maximum 200 Words)</b>  The Information Assurance Technical Framework (IATF) document was developed to help a broad audience of users both define and understand their technical needs as well as select approaches to meet those needs. The intended audience includes system security engineers, customers, scientists, researchers, product and service vendors, standards bodies, and consortia. The objectives of the IATF include raising the awareness of IA technologies, presenting the IA needs of Information System (IS) users, providing guidance for solving IA issues, and high-lighting gaps between current IA capabilities and needs. Chapter 1 outlines the information infrastructure, the information infrastructure boundaries, the information assurance framework areas, and general classes of threats. Then the Defense-in-Depth Strategy is introduced and the overall organization of the IATF document is presented.				
<b>14. SUBJECT TERMS</b> IATAC Collection, Information Assurance Technical Framework (IATF), Network, Infrastructure, ISSE, Countermeasures,			<b>15. NUMBER OF PAGES</b>	
			<b>16. PRICE CODE</b>	
<b>17. SECURITY CLASSIFICATION OF REPORT</b> UNCLASSIFIED	<b>18. SECURITY CLASSIFICATION OF THIS PAGE</b> UNCLASSIFIED	<b>19. SECURITY CLASSIFICATION OF ABSTRACT</b> UNCLASSIFIED	<b>20. LIMITATION OF ABSTRACT</b>  UNLIMITED	

***Please review and provide comments.***

The Information Assurance Technical Framework is an evolving document. It will be expanded and updated. For these changes to be most beneficial, **your** comments and suggestions are needed. Please provide any comments or suggestions you care to make to:

IATF Forum Webmaster  
C/O: Booz·Allen & Hamilton Inc.  
900 Elkridge Landing Road (Airport Square 2)  
Linthicum, MD 21090

Telephone: (410) 684-6246  
Fax: (410) 684-6487  
E-mail: [webmaster@iatf.net](mailto:webmaster@iatf.net)



# Foreword

The Information Assurance Technical Framework (IATF) document, Release 3.0, provides technical guidance for protecting United States (U.S.) Government and industry information and information infrastructures. Today, the information infrastructure processes, stores, and transmits information critical to the mission/business operations of the organization. Protecting this information is achieved through “Information Assurance” (IA) that addresses the full suite of security requirements for today's information infrastructure. Information assurance relies on ***the people, the operations, and the technology*** to accomplish the mission/business and to manage the technology/information infrastructure. Attaining a robust information assurance posture means implementing policies, procedures, techniques, and mechanisms at all layers throughout the organization's information infrastructure.

The IATF defines a process for developing a system with information assurance and the security requirements for the hardware and software components in the system. Applying these principles results in layers of protection in the information infrastructure known as the Defense-in-Depth Strategy. The four major technology focus areas of the Defense-in-Depth Strategy are Defend the Network and Infrastructure, Defend the Enclave Boundary, Defend the Computing Environment, and Supporting Infrastructures. The Defense-in-Depth Strategy has been broadly adopted. For example, within the Department of Defense (DoD), the Global Information Grid (GIG) IA Policy and Implementation Guidance was built around the Defense-in-Depth Strategy. This departmental-level policy document ***cites*** the IATF as a source of information on technical solutions and guidance for the DoD IA implementation.

Given the evolution and broader adoption of the Defense-in-Depth Strategy for IA, the advancement of technology, and a desire to make the IATF meaningful for DoD as well as other federal government and commercial sector readers, we have created Release 3.0 of the IATF. The structure of the IATF has been updated. Older sections have been completely revised and new sections added. Throughout our development of Release 3.0 we endeavored to broaden the document beyond DoD, to “nationalize” the presentation and content. Below is a summary of the major changes from Release 2.0.1 to Release 3.0.

Completely revised Chapter 1, Introduction, and Chapter 2, Defense-in-Depth Objectives Overview, to focus directly on the Defense-in-Depth Strategy approach to IA.

Expanded Chapter 3, Information Systems Security Engineering Process, to address systems engineering, systems acquisition, risk management, certification and accreditation, and life-cycle support processes, and to show how these methodologies relate to the activities of an Information Systems Security Engineer (ISSE).

Reconfigured Chapter 4, Technical Security Countermeasures, to addresses common technical issues related to adversaries (and how they act) and to provide a discussion of the primary security services.

Expanded Chapter 6, Defend the Enclave Boundary, to address a full range of technology advances including Firewalls, Remote Access, Guards, Network Monitoring Within Enclave Boundaries and External Connections, Network Scanners Within Enclave Boundaries, Malicious Code Protection, and Multi-level Security.

## UNCLASSIFIED

Cover and Foreword  
IATF Release 3.0—September 2000

Updated Chapter 8, Supporting Infrastructure, to include a comprehensive description of Key Management Infrastructure/Public Key Infrastructure (KMI/PKI) and a discussion of Detect and Respond.

Please keep in mind that the IATF is a *living document*; the next release already is being planned. Many people provided comments and recommendations on Release 2.0.1 and their comments helped define the version of the IATF you are reading. Your suggestions, your recommendations, your needs are what will define the next release ***if we hear from you***.

We want and need your feedback.

We ask that you send us your comments, reactions, criticism, recommended changes, noted omissions, and *any suggestions* that will make this document more useful to you. Please send your suggestions to [webmaster@iatf.net](mailto:webmaster@iatf.net). We also encourage you to frequently visit the IATF Forum web site (<http://www.iatf.net>). There you will be able to see the next release of the IATF unfolding, to review examples of proposed hypertext versions of the IATF and, again, to provide us your feedback. The objective of the IATF is to be a useful document *for you*. Please let us know how we did.

On behalf of all the authors of the Information Assurance Technical Framework—Release 3.0 and its predecessors—our thanks to the many people who reviewed and commented on the IATF. Thanks also go to the many speakers and panelists of the IATF Forum sessions and the past Network Security Framework Forum sessions for sharing their valuable insights on the security architectures, standards, and solutions that industry and government are bringing to bear on the complex challenge of information assurance.

Cynthia Frederick  
IATF Technical Director

# Table Of Contents

Foreword

Executive Summary

Summary of Changes

1. Introduction
  - 1.1 Objectives
  - 1.2 Intended Audiences
  - 1.3 Context
    - 1.3.1 Information Infrastructure Defined
    - 1.3.2 Categorizing Information and Information Infrastructures
    - 1.3.3 Boundaries and Information Infrastructures
    - 1.3.4 Information Assurance Framework Areas
    - 1.3.5 Nature of Cyber Threats
  - 1.4 Defense in Depth
    - 1.4.1 Defense in Depth
  - 1.5 IATF Organization
2. Defense-in-Depth Objectives Overview
  - 2.1 Overview
    - 2.1.1 Defend the Computing Environment Objectives
    - 2.1.2 Defend the Enclave Boundary Objectives
    - 2.1.3 Defend the Network and Infrastructure Objectives
    - 2.1.4 Supporting Infrastructures Objectives
  - 2.2 Examples of User Environments
    - 2.2.1 Federal Computing Environment
    - 2.2.2 Department of Defense Computing Environment
3. Information Systems Security Engineering Process
  - 3.1 Introduction to Chapter 3
  - 3.2 Information Systems Security Engineering
    - 3.2.1 Purpose and Scope of Section 3.2
    - 3.2.2 Overview of the Systems Engineering Process as a Basis for ISSE
    - 3.2.3 The ISSE Process
  - 3.3 Relation of ISSE Process to Other Processes
    - 3.3.1 Purpose and Outline of 3.3
    - 3.3.2 Systems Acquisition Process
    - 3.3.3 Risk Management
    - 3.3.4 Life-Cycle Support
    - 3.3.5 Certification and Accreditation
    - 3.3.6 Common Criteria and Possible Uses

- 4. Technical Security Countermeasures
  - 4.1 Introduction
  - 4.2 Adversaries, Motivations, Capabilities, and Categories of Attacks
    - 4.2.1 Potential Adversaries
    - 4.2.2 Categories of Attacks
  - 4.3 Primary Security Services
    - 4.3.1 Access Control
    - 4.3.2 Confidentiality
    - 4.3.3 Integrity
    - 4.3.4 Availability
    - 4.3.5 Non-Repudiation
  - 4.4 Important Security Technologies
  - 4.5 Robustness Strategy
    - 4.5.1 Overview of the General Process
    - 4.5.2 Determining the Degree of Robustness
    - 4.5.3 Strength of Mechanism
    - 4.5.4 Level of Assurance
    - 4.5.5 A Worked Example
    - 4.5.6 Robustness Strategy Evolution
    - 4.5.7 Real World Applications
  - 4.6 Interoperability Framework
    - 4.6.1 Major Elements of Interoperability
    - 4.6.2 Challenges for Interoperability
    - 4.6.3 Interoperability Strategy
  - 4.7 Key Management Infrastructure/Public Key Infrastructure Considerations
    - 4.7.1 KMI/PKI Overview
    - 4.7.2 KMI/PKI Operational Services
    - 4.7.3 KMI/PKI Processes
- 5. Defend the Network and Infrastructure
  - 5.1 Availability of Backbone Networks (BNs)
    - 5.1.1 Target Environment
    - 5.1.2 Consolidated Requirements
    - 5.1.3 Potential Attacks and Potential Countermeasures
    - 5.1.4 Technology Assessment
    - 5.1.5 Framework Guidance
  - 5.2 Wireless Networks Security Framework
    - 5.2.1 Cellular Telephone
    - 5.2.2 LEO/MEO Satellite Telephone Networks
    - 5.2.3 Wireless Local Area Network
    - 5.2.4 Paging (One-way and Two-way)
    - 5.2.5 WLL/Wireless PBX/Cordless Telephones
  - 5.3 System High Interconnections and Virtual Private Networks (VPNs)
    - 5.3.1 Target Environment
    - 5.3.2 Consolidated Requirements
    - 5.3.3 Potential Attacks
    - 5.3.4 Potential Countermeasures

- 5.3.5 Technology Assessment
- 5.3.6 Cases
- 5.3.7 Framework Guidance
- 5.4 Secure Voice
- 5.5 Multiple Security Layers
- 6. Defend the Enclave Boundary/External Connections
  - 6.1 Firewalls
    - 6.1.1 Target Environment
    - 6.1.2 Firewall Requirements
    - 6.1.3 Potential Attacks
    - 6.1.4 Potential Countermeasures
    - 6.1.5 Firewall Technology Assessment
    - 6.1.6 Cases
    - 6.1.7 Enclave Boundary Protection Framework Guidance
  - 6.2 Remote Access
    - 6.2.1 Target Environment
    - 6.2.2 Consolidated Requirements
    - 6.2.3 Potential Attacks
    - 6.2.4 Potential Countermeasures
    - 6.2.5 Technology Assessment
    - 6.2.6 Cases
    - 6.2.7 Framework Guidance
  - 6.3 Guards
    - 6.3.1 Target Environment
    - 6.3.2 Requirements
    - 6.3.3 Potential Attacks
    - 6.3.4 Potential Countermeasures
    - 6.3.5 Guard Technology Assessment
    - 6.3.6 Selection Criteria
    - 6.3.7 Framework Guidance
    - 6.3.8 Technology Gaps
  - 6.4 Network Monitoring Within Enclave Boundaries and External Connections
    - 6.4.1 Network Intrusion Detection
    - 6.4.2 Malicious Code (or Virus) Detectors
    - 6.4.3 Discussion of Typical Bundling of Capabilities
    - 6.4.4 Beyond Technology Solutions
    - 6.4.5 For More Information
  - 6.5 Network Scanners Within Enclave Boundaries
    - 6.5.1 Network Vulnerability Scanners
    - 6.5.2 War Dialers
    - 6.5.3 Considerations for Deployment
    - 6.5.4 Considerations for Operation
    - 6.5.5 Beyond Technology Solutions
    - 6.5.6 For More Information
  - 6.5 Malicious Code Protection
    - 6.6.1 Target Environment

- 6.6.2 Malicious Code Protection Requirements
  - 6.6.3 Potential Attack Mechanisms
  - 6.6.4 Potential Countermeasures
  - 6.6.5 Technology Assessment
  - 6.6.6 Selection Criteria
  - 6.6.7 Cases
  - 6.6.8 Framework Guidance
- 6.7 Multi-Level Security (MLS)
  - 6.7.1 High-to-Low
  - 6.7.2 MLS Workstation
  - 6.7.3 MLS Servers
  - 6.7.4 MLS Network Components
- 7. Defend the Computing Environment
  - 7.1 Security for System Applications
    - 7.1.1 Target Environment
    - 7.1.2 Consolidated Requirements
    - 7.1.3 Potential Attacks
    - 7.1.4 Potential Countermeasures
    - 7.1.5 Technology Assessment
    - 7.1.6 Cases
    - 7.1.7 Framework Guidance
  - 7.2 Host-Based Detect and Respond Capabilities Within Computing Environments
    - 7.2.1 Host Monitors—Intrusion Detection
    - 7.2.2 Host Monitors— Malicious Code or Virus Detectors
    - 7.2.3 Host Scanners—Host Vulnerability Scanners
    - 7.2.4 Host Scanners—File Integrity Checkers
    - 7.2.5 Discussion of Typical Bundling of Capabilities Within Products
    - 7.2.6 Beyond Technology Solutions
    - 7.2.7 For More Information
- 8. Supporting Infrastructure
  - 8.1 Key Management Infrastructure/ Public Key Infrastructure (KMI/PKI)
    - 8.1.1 KMI/PKI Introduction
    - 8.1.2 Certificate Management
    - 8.1.3 Symmetric Key Management
    - 8.1.4 Directories
    - 8.1.5 Infrastructure Management
    - 8.1.6 KMI/PKI Assurance
    - 8.1.7 KMI/PKI Solutions
    - 8.1.8 FUTURE TRENDS of Public Key Infrastructure
  - 8.2 Detect and Respond as a Supporting Element
    - 8.2.1 What This Focus Area Addresses
    - 8.2.2 Enterprise Architecture Considerations
    - 8.2.3 General Considerations for a Detect and Respond Solution
    - 8.2.4 Detect and Respond Functions
    - 8.2.5 Relevant Detect and Respond Technologies

8.2.6 For More Information

- 9. Information Assurance for the Tactical Environment
  - 9.1 Target Environment
  - 9.2 Wiping Classified Data from Tactical Equipment
    - 9.2.1 Mission Need
    - 9.2.2 Consolidated Requirements
    - 9.2.3 Technology Assessment
    - 9.2.4 Framework Guidance
  - 9.3 Stored Data Protection in a Hostile Environment
    - 9.3.1 Mission Need
    - 9.3.2 Consolidated Requirements
    - 9.3.3 Technology Assessment
    - 9.3.4 Framework Guidance
  - 9.4 Key Management in a Tactical Environment
    - 9.4.1 Mission Need
    - 9.4.2 Consolidated Requirements
    - 9.4.3 Technology Assessment
    - 9.4.4 Framework Guidance
  - 9.5 Network Mobility/Dynamic Networks
    - 9.5.1 Mission Need
    - 9.5.2 Consolidated Requirements
    - 9.5.3 Technology Assessment
    - 9.5.4 Framework Guidance
  - 9.6 Access to Individual Classified Accounts by Multiple Users
    - 9.6.1 Mission Need
    - 9.6.2 Consolidated Requirements
    - 9.6.3 Technology Assessment
    - 9.6.4 Framework Guidance
  - 9.7 Secure Net Broadcast/Multicast
    - 9.7.1 Mission Need
    - 9.7.2 Consolidated Requirements
    - 9.7.3 Technology Assessment
    - 9.7.4 Framework Guidance
  - 9.8 IA Solutions in Low Bandwidth Communications
    - 9.8.1 Mission Need
    - 9.8.2 Consolidated Requirements
    - 9.8.3 Technology Assessment
    - 9.8.4 Framework Guidance
  - 9.9 Split-Base Operations
    - 9.9.1 Mission Need
    - 9.9.2 Consolidated Requirements
    - 9.9.3 Technology Assessment
    - 9.9.4 Framework Guidance
  - 9.10 Multi-Level Security (MLS)
    - 9.10.1 Mission Need
    - 9.10.2 Consolidated Requirements

- 9.10.3 Technology Assessment
- 9.10.4 Framework Guidance
- 9.11 Additional Technologies

## 10. A View of Aggregated Solutions

### APPENDICES

- A. Acronyms
- B. Glossary
- C. Characterization of Customer Community Networks
- D. System Security Administration (SSA)
- E. Office of the Secretary of Defense (OSD) Information Assurance (IA) Policy  
Robustness Levels
- F. Executive Summaries
- G. Protection Profiles



# List of Figures

## Chapter 1

Figure 1-1	Availability & Protection Requirements
Figure 1-2	Information Infrastructure Elements
Figure 1-3	IA Technology Framework Areas
Figure 1-4	Local Computing Environment Area
Figure 1-5	Enclave Boundaries Framework Area
Figure 1-6	Network and Infrastructure Framework Area
Figure 1-7	Classes of Attacks on the Information Infrastructure
Figure 1-8	Principal Aspects of the Defense-in-Depth Strategy
Figure 1-9	Composition of the IATF

## Chapter 2

Figure 2-1	Example Federal Computing Environment—DOE
Figure 2-2	Example Federal Computing Environment—DoD

## Chapter 3

Figure 3-1	Systems Engineering Process
Figure 3-2	Impact of Mission, Threats, and Policies on Information Protection Requirements
Figure 3-3	Layered Requirements “Hierarchy”
Figure 3-4	ISSE and System Acquisition Process Flows
Figure 3-5	Risk Management Process
Figure 3-6	Risk Decision Flow
Figure 3-7	Risk Plane
Figure 3-8	DITSCAP Flow
Figure 3-9	Security Concepts and Relationships in the Common Criteria
Figure 3-10	Protection Profile — Security Target Relationship
Figure 3-11	Evaluation Concepts and Relationships
Figure 3-12	Uses of Evaluation Results

## Chapter 4

Figure 4-1	Categories of Attacks Against Networked Systems
------------	---

## Chapter 5

Figure 5-1	Defend the Network & Infrastructure
Figure 5.1-1	Backbone Availability Model
Figure 5.2-1	Wireless Extension of the Wired Infrastructure
Figure 5.2-2	Cellular Telephone Environment

Figure 5.2-3	Mobile Satellite Subscriber Environment
Figure 5.2-4	WLAN Environment
Figure 5.2-5	Pager Environment
Figure 5.2-6	Pager Usage Scenario
Figure 5.2-7	Wireless Telephony Environments
Figure 5.3-1	Target Environment Communications Infrastructure
Figure 5.3-2	Local Virtual Private Network Architectures
Figure 5.3-3	IP Layering Encryption Methods
Figure 5.3-4	Reverse Tunneling Placement of Cryptographic Mechanisms

## **Chapter 6**

Figure 6-1	Defend the Enclave Boundary
Figure 6.1-1	Enclave Boundary Environment
Figure 6.1-2	Application Gateway
Figure 6.1-3	Dual-Homed Firewall Architecture
Figure 6.1-4	Screened Host Firewall Architecture
Figure 6.1-5	Screened Subnet Firewall Architecture
Figure 6.1-6	Case 1: Private to Public Network Communication
Figure 6.1-7	Case 2: Remotely Accessing a Private Network
Figure 6.1-8	Case 3: Private Network Connectivity via a Lower-Level Network
Figure 6.1-9	Case 4: Collaborative LAN's with Public Network Connections
Figure 6.2-1	Typical Remote Access Environment
Figure 6.2-2	Attacks Against the Remote Access Scenario
Figure 6.2-3	Security Technologies in the Remote Access Scenario
Figure 6.2-4	Protocol Layers In Remote Access Scenario
Figure 6.2-5	Remote Access Cases
Figure 6.3-1	Guard Environment
Figure 6.3-2	Dual Network Approach
Figure 6.3-3	Cascading Protection
Figure 6.3-4	File Transfers
Figure 6.3-5	Secret to SBU Releasability
Figure 6.3-6	Human Reviewer-Man in the Middle
Figure 6.3-7	Releasability Human Verification
Figure 6.3-8	Sanitization Process
Figure 6.4-1	Breakdown of Network Monitor Technologies
Figure 6.4-2	Network IDS Deployment Options
Figure 6.5-1	Back door Attacks Through Telephone Networks
Figure 6.6-1	Malicious Code Relationships
Figure 6.6-2	Sources of Malicious Code Infections
Figure 6.6-3	Virus Execution
Figure 6.6-4	Logic Bomb Execution

Figure 6.6-5	Virus Filter
Figure 6.6-6	DOS File Infection
Figure 6.6-7	Intelligent Scanning Architecture (ISA)
Figure 6.6-8	Macro Virus Infection
Figure 6.6-9	Polymorphic Virus Infection
Figure 6.6-10	Trojan Horse Infection
Figure 6.7-1	High To Low Concepts
Figure 6.7-2	Recommended Topology

## Chapter 7

Figure 7-1	Local Computing Environment
Figure 7.1-1	Custom N-Tier Application
Figure 7.2-1	Breakdown of Host Sensor Technologies

## Chapter 8

Figure 8-1	Supporting Infrastructures: KMI/PKI
Figure 8-2	Supporting Infrastructures: Detect and Respond
Figure 8.1-1	Interactions of the KMI/PKI Applications Operational Services
Figure 8.1-2	Using Public Key Infrastructures in Secure Enclaves
Figure 8.1-3	Hierarchical, Trust List and Mesh Approaches to PKI Interoperation
Figure 8.1-4	Bilateral Cross-Certification, Bridge CA, and on-Line Status Approaches to PKI Interoperation
Figure 8.1-5	Browser Certification: Key Generation and Certificate Request
Figure 8.1-6	Browser Certification: CA Processing Request
Figure 8.1-7	S/MIME Client Certification Process
Figure 8.1-8	Browser Certification: Installing Certificate in Browser
Figure 8.1-9	Critical Elements of Symmetric Key Management Activities
Figure 8.1-10	The Directory Model
Figure 8.1-11	Directory Use Access
Figure 8.1-12	Example of the Key Management Infrastructure Directory Information Tree
Figure 8.1-13	Access Control Decision Function Required for Access Control
Figure 8.1-14	DoD Class 3 PKI Architecture
Figure 8.1-15	FORTEZZA CMI Components
Figure 8.1-16	Operational Activities Supported by the K MI
Figure 8.1-17	DoD KMI System Context
Figure 8.1-18	Nodal View of the Target KMI
Figure 8.1-19	Breakdown of Client Nodes
Figure 8.1-20	Federal PKI Architecture
Figure 8.2-1	Perspectives of Layers in a Detect and Respond Infrastructure Hierarchy
Figure 8.2-2	Basic Hierarchy for Detect and Respond Infrastructure
Figure 8.2-3	Basic View of Detect and Respond Phases

List of Figures

IATF Release 3.0—September 2000

Figure 8.2-4	Realistic View of Detect and Respond Phases
Figure 8.2-5	Possible Allocations of Detect and Respond Functions
Figure 8.2-6	Functions to Support Warning
Figure 8.2-7	Functions to Support Local Incident Detection
Figure 8.2-8	Functions to Support Incident Characterization
Figure 8.2-9	Functions to Support Incident Response
Figure 8.2-10	Functions to Support Attack Determination
Figure 8.2-11	Functions to Support Attack Characterization
Figure 8.2-12	Functions to Support Response Coordination
Figure 8.2-13	Functions to Support Attack Investigation
Figure 8.2-14	Detect and Respond Technologies
Figure 8.2-15	Sensor Technologies Grouping
Figure 8.2-16	Possible Sensor Deployment Locations
Figure 8.2-17	Detect and Respond Technology Reference Model

## Chapter 9

Figure 9-1	The Tactical Communications Environment
Figure 9-2	Tactical Communications Information Flow
Figure 9-3	Interconnecting Cell Sites using a UA V
Figure 9-4	Near-Term Architecture
Figure 9-5	Objective WIN Security Architecture
Figure 9-6	Battlefield Video Teleconference

# List of Tables

## Chapter 1

Table 1-1     Classes of Attack

## Chapter 4

Table 4-1	Examples of Potential Adversaries
Table 4-2	Examples of Specific Passive Attacks
Table 4-3	Examples of Active Attacks
Table 4-4	Examples of Close-In Attacks
Table 4-5	Examples of Insider Attacks
Table 4-6	Examples of Distribution Attacks
Table 4-7	Degree of Robustness
Table 4-8	Security Management Mechanisms
Table 4-9	Confidentiality Mechanisms
Table 4-10	Integrity Mechanisms
Table 4-11	Availability Mechanisms
Table 4-12	Identification and Authentication Mechanisms
Table 4-13	Access Control Mechanisms
Table 4-14	Accountability Mechanisms
Table 4-15	Non-Repudiation Mechanisms
Table 4-16	Example Depicting Use of Degree of Robustness Table
Table 4-17	Example Depicting Use of Confidentiality Mechanisms Table
Table 4-18	Example Depicting Use of Security Management Mechanisms Table
Table 4-19	Example Depicting Use of Degree of Robustness Table

## Chapter 5

Table 5.3-1	Digital Service Standards
Table 5.3-2	Characteristics of Layer 2 Protected Networks
Table 5.3-3	Characteristics of Layer 3 Protected Networks

## Chapter 6

Table 6.2-1a	Summary Guidance for Remote Access Direct Dial-up Access to Secret Enclave
Table 6.2-1b	Summary Guidance for Remote Access Direct Dial-up Access to Secret Enclave
Table 6.2-1c	Summary Guidance for Remote Access Direct Dial-Up Access to Secret Enclave
Table 6.2-1d	Summary Guidance for Remote Access Direct Dial-up Access to Secret Enclave
Table 6.4-1	Network-based IDS Considerations
Table 6.6-1	Comparison of Macro Viruses

## Chapter 7

Table 7.2-1 Host-Based IDS Considerations

Table 7.2-2 File Integrity Checker Considerations

## Chapter 8

Table 8.1-1 KMI/PKI Services Support to Subscriber Categories

Table 8.1-2 Security Applications Supported By Cryptographic Type

Table 8.1-3 KMI/PKI Processes

Table 8.1-4 Attacks and Countermeasures

Table 8.1-5 Business Requirement and Security Technology Comparison

## Appendix C

Table C-1 Public/Commercial Networks (Satellites)

Table C-2 Public/Commercial Networks

Table C-3 DoD Networks

Table C-4 Networking Technologies

# Errata Sheet

## Known Errors In the IATF Document Release 3.0

As of 01 February

IATF Chapter/ Section	What's Present in Release 3.0	Correct Information	Updated
Chapter 4.1	Old, Cross-reference: Chapter 3, Information Systems Security Methodology	Updated Cross-reference Is: Chapter 3, Information systems Security Engineering Process	January 30, 2001
Chapter 4.3.2.1	Old, Cross-reference: Chapter 3, Information Systems Security Methodology	Updated Cross-reference Is: Chapter 4.2.2.2 Active Attacks	January 30, 2001
Chapter 6.1.6 and Reference	Old, No Longer Valid Web Address: <a href="http://www.armadillo.huntsville.al.us/">http://www.armadillo.huntsville.al.us/</a>	Updated Web Address Is: <a href="http://www.fortezza-support.com">http://www.fortezza-support.com</a>	January 30, 2001
Chapter 6.3 Reference	Old, No Longer Valid Web Address: <a href="http://www.Europe.F-Secure.com/solutions/white-papers/ipsec0399.htm">http://www.Europe.F-Secure.com/ solutions/white-papers/ipsec0399.htm</a>	Updated Web Address Is: <a href="http://www.cisco.com/warp/public/cc/techno/protocol/ipsecur/ipsec/tech/ipsec_wp.htm">http://www.cisco.com/warp/public/cc/techno/ protocol/ipsecur/ipsec/tech/ipsec_wp.htm</a>	January 30, 2001
Chapter 6.5.6.4	Old, No Longer Valid Web Address: <a href="http://www.seclab.cs.ucdavis.edu/cidf">www.seclab.cs.ucdavis.edu/cidf</a>	Updated Web Address Is: <a href="http://www.gidos.org/">http://www.gidos.org/</a>	January 30, 2001
Chapter 6.7.1.3	Old, Cross-reference: Chapter 3, System Security Methodology	Updated Cross-reference Is: Chapter 4.2.2. Categories of Attacks	January 30, 2001
Chapter 7.1.1.5	Old, Cross-reference: Chapter 3, System Security Methodology	Updated Cross-reference Is: Chapter 3, Information systems Security Engineering Process	January 30, 2001

8.1.7.6 and Reference	Old, No Longer Valid Web Address: <a href="http://nii.isi.edu/info/kerberos">http://nii.isi.edu/info/kerberos</a>	Updated Web Address Is: <a href="http://www.isi.edu/gost/info/kerberos/">http://www.isi.edu/gost/info/kerberos/</a>	January 30, 2001
Chapter 8.1 Reference	Old, No Longer Valid Web Address: <a href="http://nii.isi.edu/publications/kerberos-neuman-tso.html">http://nii.isi.edu/publications/kerberos-neuman-tso.html</a>	Updated Web Address Is: <a href="http://www.isi.edu/gost/publications/kerberos-neuman-tso.html">http://www.isi.edu/gost/publications/kerberos-neuman-tso.html</a>	January 30, 2001
Chapter 9.7.3 and Reference	Old, No Longer Valid Web Address: <a href="http://milsatcom.monmouth.army.mil/gbs.htm">http://milsatcom.monmouth.army.mil/gbs.htm</a>	Updated Web Address Is: <a href="http://www.monmouth.army.mil/smc/dcats/projects/dscstrm.htm">http://www.monmouth.army.mil/smc/dcats/projects/dscstrm.htm</a>	January 30, 2001



# Executive Summary

## Chapter 1 – Introduction

The Information Assurance Technical Framework (IATF) document was developed to help a broad audience of users both define and understand their technical needs as well as select approaches to meet those needs. The intended audience includes system security engineers, customers, scientists, researchers, product and service vendors, standards bodies, and consortia. The objectives of the IATF include raising the awareness of IA technologies, presenting the IA needs of Information System (IS) users, providing guidance for solving IA issues, and highlighting gaps between current IA capabilities and needs. Chapter 1 outlines the information infrastructure, the information infrastructure boundaries, the information assurance framework areas, and general classes of threats. Then the Defense-in-Depth Strategy is introduced and the overall organization of the IATF document is presented.

## Chapter 2 – Defense-in-Depth Overview

When developing an effective IA posture, all three components of the Defense-In-Depth Strategy—people, technology, and operations—need to be addressed. This framework document focuses primarily on the technology aspects of Defense-in-Depth. The technology objectives and approaches in four focus areas, explained in the sections that follow, address the needs of the private, public, civil, and military sectors of our society.

Chapter 2 provides an overview of the Defense-in-Depth technology objectives and gives two examples of federal computing environments. The Defense-in-Depth objectives are organized around the four Defense-in-Depth technology focus areas:

- Defend the Network and Infrastructure
  - Availability of Backbone Networks
  - Wireless Networks Security Framework
  - System High Interconnections and Virtual Private Networks
- Defend the Enclave Boundary
  - Protection for Network Access
  - Remote Access
  - Multi-Level Security
- Defend the Computing Environment
  - End User Environment
  - Security for System Applications
- Supporting Infrastructures
  - Key Management Infrastructure/Public Key Infrastructure (KMI/PKI)
  - Detect and Respond

## Chapter 3 – Information Systems Security Engineering Process

Chapter 3 describes the systems engineering and Information Systems Security Engineering (ISSE) processes. The ISSE process is presented as a natural extension of the systems engineering process. Both processes share common elements: discovering needs, defining system functionality, designing system elements, producing and installing the system, and assessing the effectiveness of the system. Other systems processes: systems acquisition, risk management, certification and accreditation, and life-cycle support processes – are explained in relation to the ISSE process. Chapter 3 also provides suggestions on how the Common Criteria might be used to support the ISSE process. The processes described in this chapter provide the basis for the background information, technology assessments, and guidance contained in the remainder of the IATF document.

## Chapter 4 – Technical Security Countermeasures

This chapter of the IATF provides the background for detailed technical discussions contained in later sections of the IATF. It presents a general discussion of the principles for determining appropriate technical security countermeasures. The chapter includes a detailed description of threats, including attacker motivations, information security services, and appropriate security technologies. Using the methodology described in Chapter 3 (Information Systems Security Engineering Process), threats to the information infrastructure result in the identification of vulnerabilities followed by a managed approach to mitigating risks. Chapter 4 explains how primary security mechanisms, the robustness strategy, interoperability, and Key Management Infrastructure/Public Key Infrastructure should be considered in the selection of security countermeasures, technology, and mechanisms. These decisions form the basis for developing appropriate technical countermeasures for the identified threats, based on the value of the information.

## Chapter 5 – Defend the Network and Infrastructure

Chapter 5 describes the Defend the Network and Infrastructure technology focus area of the Defense-in-Depth Strategy. The chapter describes the types of network traffic—user, control, and management—and the basic requirements to ensure that network services remain both available and secure. Organizations that operate networks should defend their networks and the infrastructures that support their networks by establishing clear Service Level Agreements (SLA) with their commercial carriers that specify metrics for reliability, priority, and access control. Organizations must recognize that their data may be unprotected during transmission and take additional steps. Chapter 5 describes current strategies for defending networks (including data, voice, and wireless) and the corresponding network infrastructures.

## Chapter 6 – Defend the Enclave Boundary/ External Connections

Defense of the enclave boundary in Chapter 6 focuses on effective control and monitoring of the data flows into and out of the enclave. Effective control measures include firewalls, guards, Virtual Private Networks (VPNs), and Identification and Authentication (I&A)/access control for remote users. Effective monitoring mechanisms include network-based Intrusion Detection System (IDS), vulnerability scanners, and virus detectors located on the Local Area Network (LAN). These mechanisms work alone, as well as in concert with each other to provide defenses for those systems within the enclave. Although the primary focus of boundary protection is on protecting the inside from the outside, protected enclave boundaries also use technology and mechanisms to protect against malicious insiders who use the enclave to launch attacks or who facilitate outsiders gaining access through open doors or covert channels. The technologies discussed in Chapter 6 include firewalls, guards, virus/malicious code detection systems, IDS, and multi-level security systems. The IA strategy for defending an enclave boundary should flexibly implement those policies governing communications both between secure enclaves and between secure enclaves and external systems. The IA strategy must also provide the management capabilities for verifying compliance with policies governing defense of the enclave boundary.

## Chapter 7 – Defend the Computing Environment

Chapter 7 discusses the third technology focus area of the Defense-in-Depth Strategy, Defend the Computing Environment. The computing environment includes the end user workstation—both desktop and laptop—including peripheral devices. Servers include application, network, web, file, and internal communication servers. A fundamental tenet of the Defense-in-Depth Strategy is prevention of cyber attacks from penetrating networks and compromising the confidentiality, integrity, and availability of the computing environment information. For those attacks that do succeed, key are early detection and effective response to mitigate the effects of attacks. Intrusion detection, network scanning, and host scanning are the measurement functions that, on a continuous or periodic basis, determine the effectiveness of the deployed protection systems. Chapter 7 also addresses host-based sensors including those that operate in near real time as well as those that operate off-line.

## Chapter 8 – Supporting Infrastructures

Supporting Infrastructures is the fourth technology focus area of the Defense-in-Depth Strategy. The IATF addresses two supporting infrastructure entities: KMI/PKI and Detect and Respond. KMI/PKI focuses on the technologies, services, and processes used to manage public key certificates and symmetric cryptography. The discussion concludes with recommendations for the features needed to achieve the three Global Information Grid defined assurance levels: basic, medium, and high. The Detect and Respond section of Chapter 8 addresses providing warnings, detecting and characterizing suspected cyber attacks, coordinating effective responses, and performing investigative analyses of attacks.

## Chapter 9 – Information Assurance for the Tactical Environment

The tactical environment, in which military or military-style operations are conducted, presents unique information assurance challenges. In this operational environment, there is heavy reliance on the communication of urgent, time-sensitive, or life-and-death information often over wireless links. . In the past, tactical communications equipment was primarily composed of Government-Off-The-Shelf (GOTS) equipment. Decreased budgets and increased interoperability requirements in today's military organizations have led to the increased use of commercially developed equipment in tactical communications. Included in this use of commercial equipment is the use of commercial wireless networks and equipment in the tactical environment. Chapter 9 discusses the information assurance needs of the tactical environment, highlighting key tactical issues and identifying the associated security implications.

## Chapter 10 – A View of Aggregated Solutions

This section of the Framework is included in recognition of the fact that the needs of most users are represented not by any single technology focus area, but by some combinations of them. A future release of the Framework will include a discussion of developing and evaluating security approaches that are aggregations of the recommendations from the individual categories.

## In Closing ...

This Framework document is principally intended as a reference document to provide insight and guidance to security managers and system security engineers into how to address the information assurance concerns of their organizations. It is tutorial (vice prescriptive) in nature in recognition of the fact that many organizations face unique challenges that don't lend themselves to "one size fits all" solutions. This document offers insights intended to help improve the community awareness of the tradeoffs among available solutions (at a technology, not product level) and of the desired characteristics of information assurance approaches for particular problems. While this Framework attempts to lay out a large amount of information in an orderly sequence, it is structured to allow readers to use the table of contents to find topics of interest.

# Summary of Changes

As of September 2000

This section summarizes the changes that have been made to the Framework document with each release, beginning with today's IATF Release 3.0 through the initial draft NSF documents.

In general, with each release spelling errors are corrected; editing, formatting, and punctuation changes are made. Internet URLs and acronyms are reviewed and updated as required. Framework sections are selectively updated or new sections are added. Figures are reviewed and redrawn as needed.

## Changes in IATF Release 3.0 — September 2000

- Expanded the document beyond DoD by “nationalizing” its presentation and content.
- Revised Chapter 1, Introduction and Chapter 2, Defense-in-Depth Objectives Overview to directly focus on the Defense-in-Depth Strategy approach to IA
- Expanded and renamed Chapter 3, Information Systems Security Engineering Process to address systems engineering, systems acquisition, risk management, certification and accreditation, and life-cycle support and to show how these methodologies relate to the ISSE activities.
- Reconfigured Chapter 4 to addresses the common technical issues of adversaries (and how they act) and to provide a discussion of the primary security services. Adversaries, Threat (Motivations/Capabilities), and Attacks (IATF 2.0.1, Section 3.2.2 ) became elements of Chapter 4.
- Expanded Chapter 6, Defend the Enclave Boundary/External Connections to include:
  - Added new Sections 6.4, Network Monitoring Within Enclave Boundaries and External Connections, 6.5, Network Scanners Within Enclave Boundaries, 6.6, Malicious Code Protection.
  - Revised sections 6.1, Firewall, and 6.3, Guards.
  - Moved Section 6.3, Multi-Level Security to Section 6.7.
- Added new Section 7.2, Host-Based Detect and Respond Capabilities Within Computing Environments.
- Updated Chapter 8, Supporting Infrastructure, to include both a comprehensive description of what constitutes the Key Management Infrastructure/Public Key Infrastructure (KMI/PKI) and a discussion of Detect and Respond for providing warnings, detecting and characterizing suspected cyber attacks, coordinating effective responses, and performing investigative analyses of attacks.
- Incorporated old Appendix E into Chapter 8.0 Supporting Infrastructure.
- Created new Appendix E, Office of the Secretary of Defense (OSD) Information Assurance (IA) Policy Robustness Levels.

## Changes in IATF Release 2.0.1 — 22 September 1999

Release 2.0.1 changes consisted mostly of formatting and graphical updates. These changes include:

- Redrew the remaining graphics retained from Release 1.1 for greater clarity and consistency.
- Corrected some acronyms.
- Updated table formats and headings.
- Changed the page heading to “IATF Release 2.0.1—September 1999.”—

## Changes in IATF Release 2.0 — 31 August 1999

- Name changed to Information Assurance Technical Framework (IATF).
- Alignment of the security solution frameworks with the four focus areas of the defense-in-depth strategy: Defend the Network and Infrastructure (Chapter 5), Defend the Enclave Boundary/External Connections (Chapter 6), Defend the Computing Environment (Chapter 7), and Supporting Infrastructures (Chapter 8).
- System High Interconnections and Virtual Private Networks (VPNs) (NSF-R1.1 Section 5.2) and Availability of Backbone Networks (NSF R1.1 Section 5.7) became elements of the new Chapter 5, Defend the Network and Infrastructure.
- Protection for Network Access (NSF R1.1 Section 5.3), Remote Access (NSF R1.1Section 5.4), and Multi-Level Security (NSF R1.1 Section 5.5) became elements of the new Chapter 6, Defend the Enclave Boundary/External Connections.
- Security for System Applications (NSF R1.1 Section 5.6) became an element of the new Chapter 7, Defend the Computing Environment.
- NSF R1.1Chapter 6 Security Management Infrastructure (SMI) was renamed Key Management Infrastructure/Public Key Infrastructure (KMI/PKI) and became an element of the new Chapter 8, Supporting Infrastructures.
- A new section, Wireless Security Solutions, was added in Chapter 5, Defend the Network and Infrastructure.
- A new Chapter 9, Information Assurance for the Tactical Environment was added.
- The outline of a new section, Detect and Respond, was added to in Chapter 8.
- Added 2 new appendices Executive Summaries (Appendix F) and Protection Profiles (Appendix G).
- Chapter 1 was revised to include an explanation of the relationship of the GNIE IA effort, the defense-in-depth strategy, and the IATF.
- The Remote Access section was updated.
- Added “UNCLASSIFIED” to the header and footer of every page.
- Redrew some of the graphics retained from Release 1.1 for greater clarity and consistency.

## Changes in NSF Release 1.1 — 3 December 1998

- A (new or updated) Robustness section for Chapter 4.
- Complete revision of Sections 5.6 (Security for System Applications) and 5.7 (Availability of Backbone Networks).
- Inclusion of Appendix A (Abbreviations & Acronyms).
- A significantly expanded Chapter 4 focused on security services, security robustness, and secure interoperability.

## Changes in NSF Release 1.0 — 22 May 1998

- Added a new Chapter 3 focused on Security Methodology.
- Added a new Chapter 4 focused on security services, security robustness, and secure interoperability.
- Added two new sections within Chapter 5 focused on security for system applications and backbone availability.
- Added a new Chapter 6 focused on Security Management Infrastructure.
- Added appendices providing a glossary of terms and amplifying information for some of the security solutions framework.
  - Glossary (Appendix B)
  - Characterization of Customer Community (Appendix C)
  - System Security Administration (Appendix D)
  - Public Key Infrastructure (PKI) Formats (Appendix E)

## The Initial Network Security Framework (NSF) Document

The first releases of the Network Security Framework (Releases 0.1 and 0.2) provided initial insight and guidance to a few categories of network security challenges. The third release (Release 1.0) provided an initial treatment of all of the primary topics that were suggested in the original outline and in the comments received.

**This page intentionally left blank**



# Chapter 1

## Introduction

The Information Assurance Technical Framework (IATF) exists to address questions such as:

- *How do I go about defining information protection needs and solutions?*
- *What technology exists to give the protection I need?*
- *What organizational resources are available to help locate the protection I need?*
- *What kind of markets exist for Information Assurance (IA) products and services?*
- *Where should research in IA approaches and technology be focused?*
- *What are the principles of IA?*

This evolving document is published to provide recommendations and information on current information assurance concerns and practices to System Security Engineers and others who address IA in their work. Over time it will reflect changes in policy, technology, environments, and the uses made of systems that depend upon information.

## 1.1 Objectives

The Framework has several objectives:

- Raise the awareness among users of information-dependent systems of information assurance technologies.
- Identify technical solutions to IA needs in accordance with national policies.
- Employ the technology focus areas of a *Defense-in-Depth* strategy to define approaches to information assurance.
- Define the security functions and protection levels needed for different situations or mission scenarios (referred to as “cases”).
- Present the IA needs of users of information-based systems.
- Highlight the need to engage a team of IA or information systems security experts to resolve pressing security needs.
- Aid the development of IA solutions that satisfy IA needs by highlighting gaps in the currently available commercial and government protection technologies.
- Provide guidance for solving IA issues by offering tutorials on available technologies, tradeoffs among available solutions (at a technology versus product level), and descriptions of desirable solutions characteristics.
- Assist purchasers of IA products by identifying important security-related features that should be sought.

## 1.2 Intended Audiences

The Framework addresses the needs of several groups of people. The following describes each group and indicates how the document can be used.

- **System security engineers:** to assist in developing IA solutions tailored to a particular customer's needs. The customer's needs can be compared with the various Framework technology areas, cases, and recommended solutions. From these, a tailored solution can be created for this particular customer.
- **Customers:** to provide answers to the myriad issues and technical challenges involved in selecting adequate IA features and assurance levels for their system and networks. Customers can include system users, managers, and security officers or administrators. With this knowledge, customers can successfully interact with security engineers and architects to design a comprehensive IA solution.
- **Scientists and researchers:** to focus their efforts on customer requirements not being met by current technology. Thus, the Framework will highlight future IA technology and identify technology gaps for use by both government and commercial research communities.
- **Commercial product and service providers:** to gain insight into the needs of customers. Industry will get an indication of the current and future markets for IA products and services.
- **Standards bodies and consortia:** to provide guidance in developing standards for commercial products. A major emphasis within the customer base focuses on the use of commercial products, which are driven by commercial standards. The IATF highlights gaps in the available standards that will help focus efforts to influence the standards bodies.

## 1.3 Context

### 1.3.1 Information Infrastructures Defined

The IATF is based on the concept of an information infrastructure. An information infrastructure comprises communications networks, computers, databases, management, applications, and consumer electronics and can exist at the global, national, or local level. The global information infrastructure is not controlled or owned by a single organization—"ownership" is distributed among corporate, academic, and government entities as well as by individuals. The Internet is an example of a global information infrastructure as is the global telecommunications network. Most organizations that communicate externally rely upon this global system in conducting their operations using a combination of global, virtual networks, dedicated networks, Wide Area Networks (WAN) and customized information systems.

A national information infrastructure is the collection of information infrastructures used by the nation to conduct its business, whether government or commercial. One instance of a national infrastructure is the United States (U.S.) critical infrastructure as defined in Presidential

Decision Directive (PDD) 63. Before the growth of multinational companies and the advent of the Internet, one could easily identify a national information infrastructure. In the last few decades however, the lines between the global and national information infrastructures have blurred significantly. Each country will need to decide whether the distinction between the two has merit; if so, criteria will be required to categorize an asset as qualifying as part of a “national” information infrastructure. In the U.S., one criterion to use might be whether assets are subject to U.S. laws, regulations, and policies.

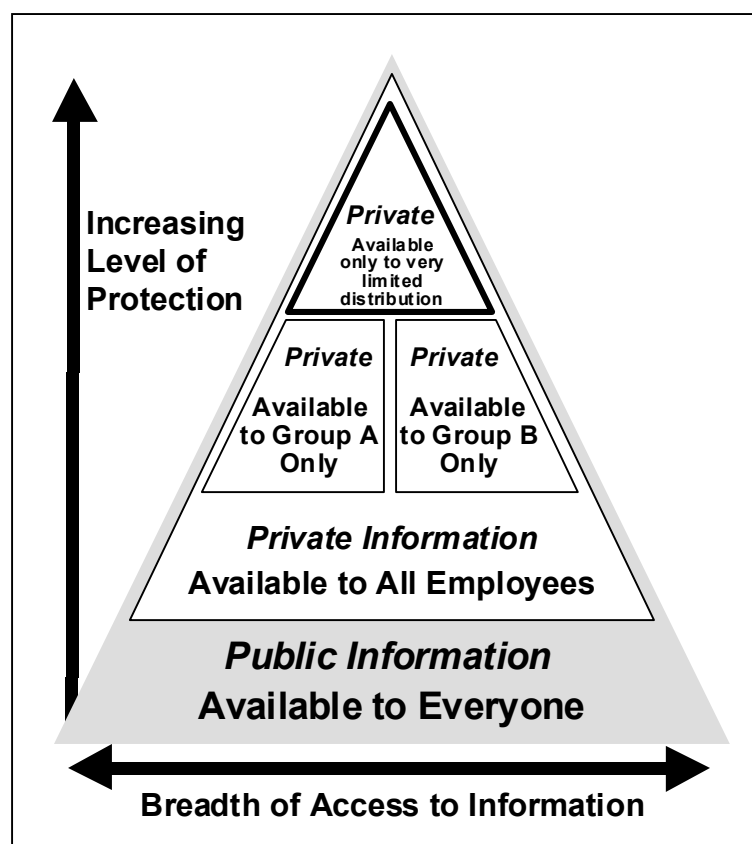
Local information infrastructures are the dedicated assets an organization operates in conducting its business; they consist mainly of commercial information systems, network technologies, and applications. Security measures are applied by the owner or operator of the local information infrastructure—defined either as an organization, or even a business unit within an organization.

### 1.3.2 Categorizing Information and Information Infrastructures

Within the organization, information processed using these assets are generally grouped into functional categories; administrative, personnel, logistics, etc. Some information may be available to the *public*, some considered *private*. There are many types of private information; companies have different types of proprietary information, government organizations have many types of classified information, including law enforcement, Secret, Top Secret, and Sensitive Compartmented Information. These divisions of information availability are also called information domains.

To accomplish their various missions and protect their critical functions, all organizations—both government and private sector—have public and private information they need to safeguard. The mission or business environment determines how and to what extent specific information is protected. What is publicly releasable to one organization may be private to another, and vice versa. The Federal Government uses specific categories for some of its private information under the heading “classified information”. In general, the Government recognizes four classification levels: unclassified, confidential, secret, and top secret. Within the classification levels, there may be subcategories specific to individual communities. Three of the classification categories—confidential, secret, and top secret—address private information. The fourth level of classification covers both some private information (such as sensitive or Privacy Act Information) and some public information.

Several types of information could be considered private. One example would be law enforcement information that could potentially damage or impair law enforcement efforts if improperly protected or handled. Proprietary information is much the same for the business community; the information would be harmful to the business if it were released. Information covered under the Privacy Act including personal financial, medical, and other such information is also considered sensitive. The Government handles a variety of classified and sensitive information supporting research, engineering, logistic, administrative and acquisition functions across the different organizations and agencies.



**Figure 1-1 Availability & Protection Requirements**

Most organizations assign more rigorous requirements to protecting their private information than their public information. First access is controlled. For example, within an organization, a human resources or finance person may have complete access to personnel and payroll databases and servers, but may not have access to the most sensitive research and development information. Within the Government—classified realm this is accomplished by assigning different classification levels, special compartments, and “need to know” designations. This is depicted in Figure 1-1.

In addition to access controls, more robust technical security measures are implemented. Organizations acknowledge that the potential loss from exposing private information to the public would be high and therefore the additional cost of

protection is warranted. In Figure 1-1, the most stringent security measures would be applied to the information and information infrastructures associated with the top triangle.

The partitioning of information according to access control, need, and levels of protection required yields categories of information. The categories are often called *information domains*. Organizations implement specific mechanisms to both enforce the information partitioning and to provide for the deliberate flow of information between information domains.

Protecting information in a collaborative environment presents its own challenges. Organizations sharing information need to agree upon the sensitivity level of the information as well as methods to protect it. Many times one organization regards information as more or less sensitive than its partner and officials from both organizations must negotiate a mutually agreeable solution. This occurs between companies sharing proprietary information, between government organizations involved in a joint project, and very often, between countries.

### 1.3.3 Boundaries and Information Infrastructures

When considering security for information infrastructures, it is important to understand the concept of boundaries. Information assets exist in physical *and* logical locations, and boundaries exist between these locations. An understanding of what is to be protected from external influences can help ensure adequate protection measures are applied where they will be most effective. However, when analyzing a real world example, this boundary is not so easily identi-

fied. Sometimes the boundary is defined as physical—people, information, and information systems associated with one physical location. But this ignores the reality that, within a single location, many different security policies may be in place, some covering public information and some covering private information.

Other times it is defined as surrounding the information and information systems that are governed by a policy within a single location. This definition, however, does not address the fact that policies cross physical boundaries. Further complicating the matter is that, many times, a single machine or server may house public *and* private information. So, multiple boundaries may exist within a single machine. Figure 1-2 illustrates these complexities associated with defining boundaries. It depicts one organization with facilities in two locations each processing multiple levels of information. In addition, the private network is also connected to the Internet. In this case, the physical location might be considered a boundary, as might the logical boundaries associated with the different levels of information.

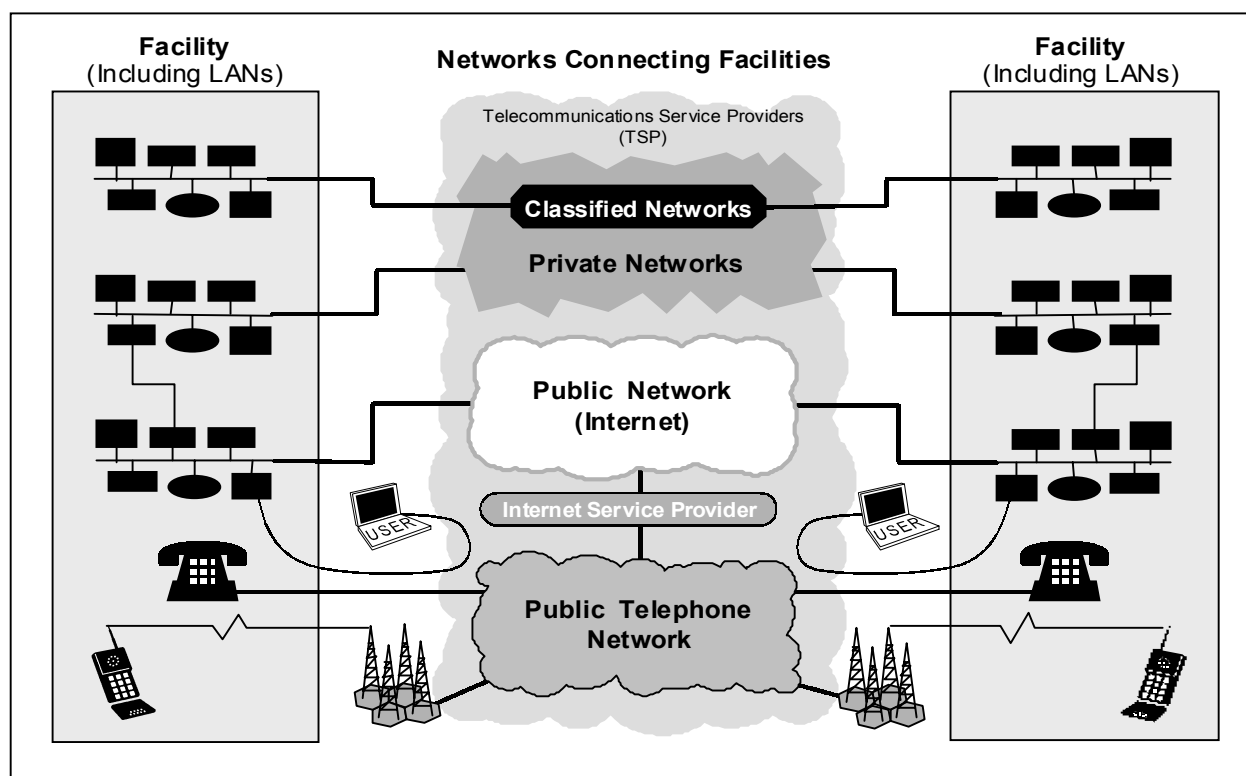


Figure 1-2 Information Infrastructure Elements

### 1.3.4 Information Assurance Framework Areas

Given the complexity of information systems, discussion of *how to protect* them is challenging unless a common framework is employed. The IATF document employs a framework which partitions the information assurance technology aspects of information systems into the following four areas, as shown in Figure 1-3.

- Local Computing Environments.
- Enclave Boundaries (around the local computing environments).
- Networks and Infrastructures.
- Supporting Infrastructures.

By partitioning the discussion by these four areas, aspects of information assurance technology

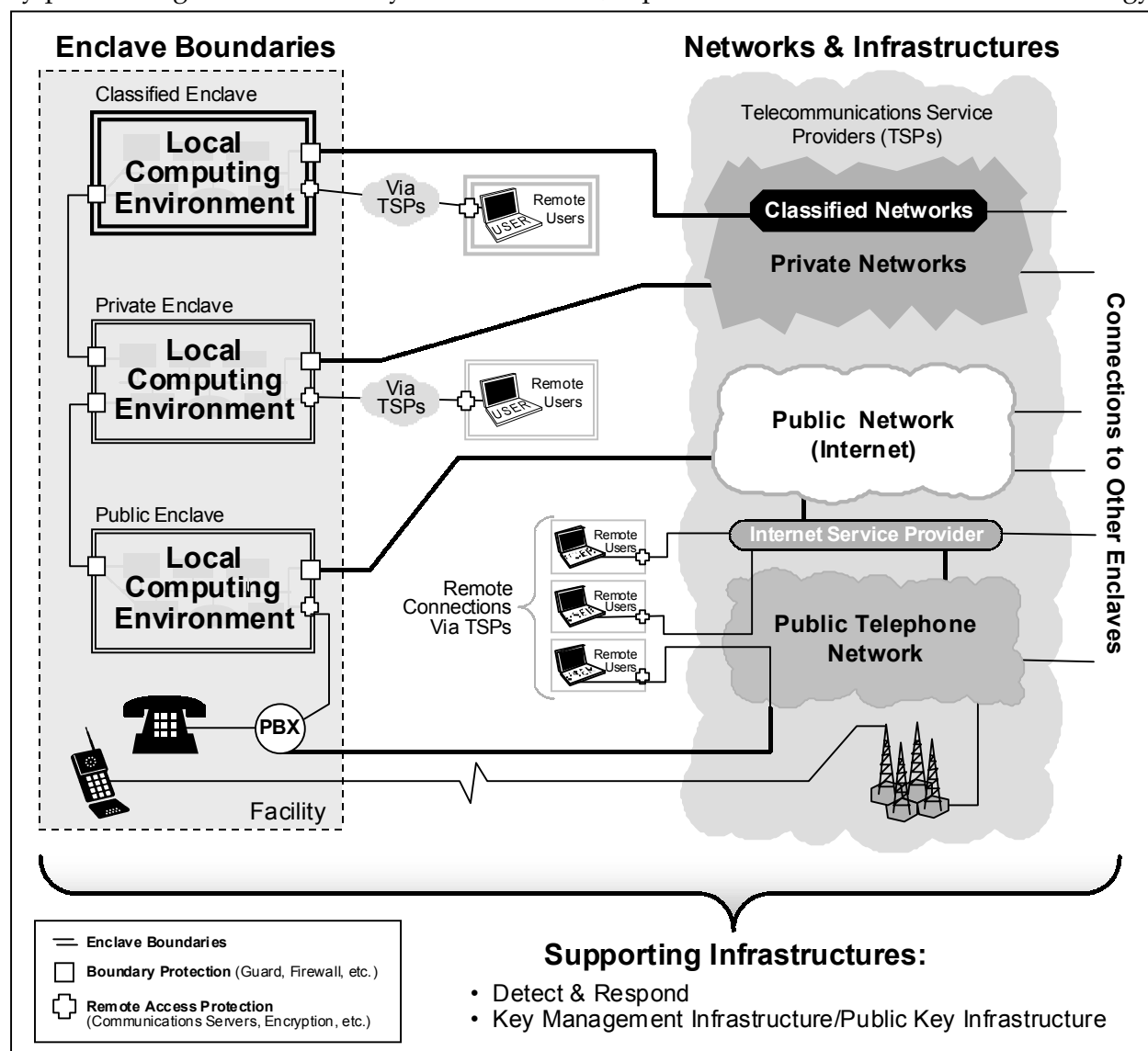


Figure 1-3 IA Technology Framework Areas

for the information system can be focused upon and more clearly presented. However, these areas are overlapping bins of concern. Effective implementation of information assurance for a given information system involves the interplay of actions taken throughout the information system—across all four technology framework areas. In the paragraphs that follow, the four framework areas are described further.

## Local Computing Environments Framework Area

The local user *computing environment* typically contains servers, clients, and the applications installed on them. Applications include, but are not limited to, those that provide services such as scheduling or time management, printing, word processing, or directories. This environment is represented in Figure 1-4.

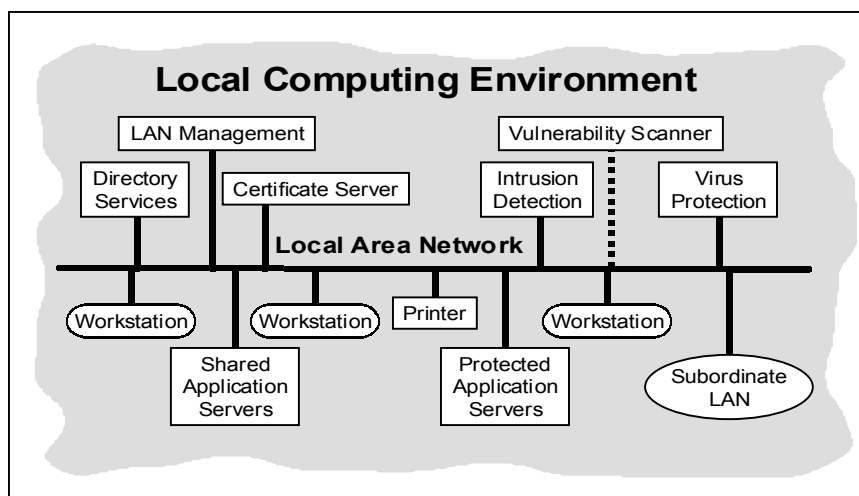
Looking across the range of computing environments, there are several broad categories of information systems that organizations employ. In both the private sector and the government, one will find large legacy information systems that have been developed over many years and at considerable expense to satisfy unique mission/business needs.

- These will likely remain in place for some time to come.

A large number of organizations have also heavily invested in the use of Commercial Off-The-Shelf (COTS) products or customized versions of COTS information system components and products tailored for their specific use. Organizations using customized products will probably transition to full COTS implementations as the product offerings address their needs more directly.

Most organizations want to use multiple applications to perform their operational mission functions. As a result users are struggling to integrate the ever-growing range of applications into an effective information processing capability. Each of these applications will place unique requirements on the Supporting Infrastructure.

Across the range of computing environments, the customer base needs IA solutions in many existing application areas. Security of the computing environment focuses on servers and clients to include the applications installed on them, the operating systems, and host-based monitoring capabilities. Application areas requiring IA solutions include the following.

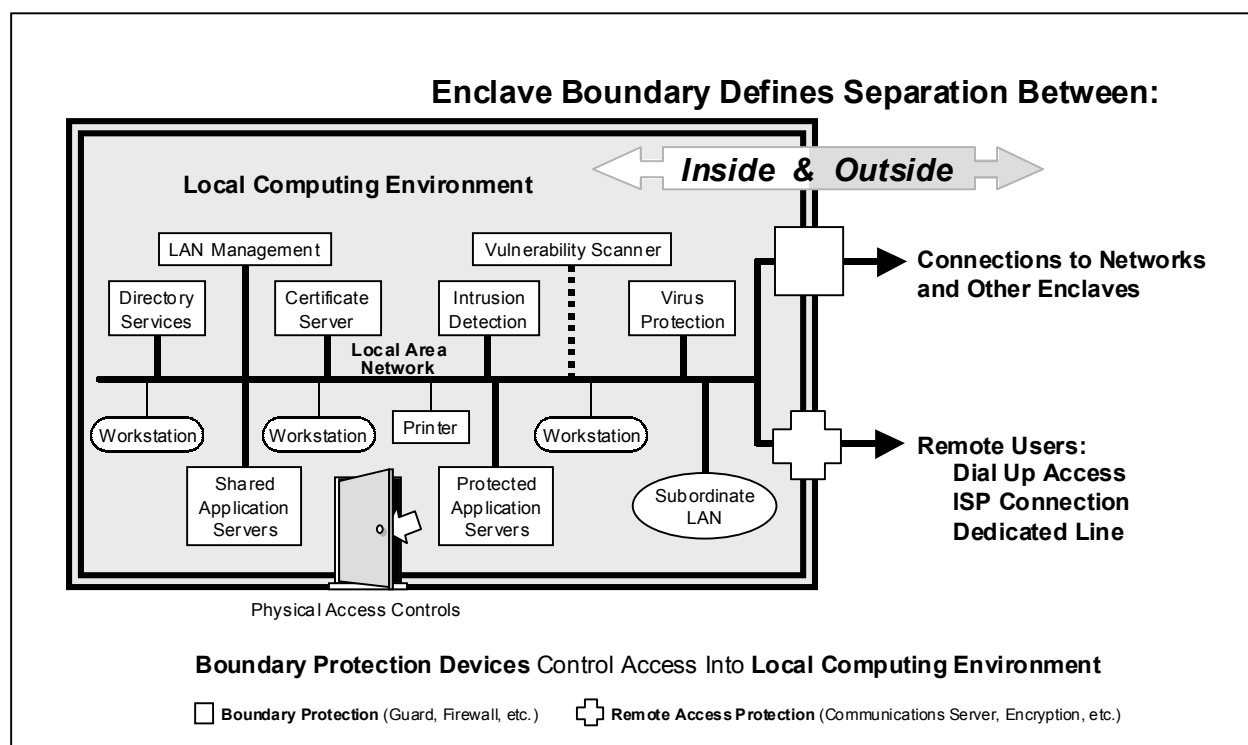


**Figure 1-4 Local Computing Environment Area**

- Messaging, e.g., electronic mail (e-mail).
- Operating systems.
- Web browser.
- Electronic commerce.
- Wireless access.
- Collaborative computing.
- Database access.

## Enclave Boundaries Framework Area

A collection of local computing devices interconnected via Local Area Networks (LAN), governed by a single security policy, regardless of physical location is considered an "enclave." As discussed above, because security policies are unique to the type, or level, of information being processed, a single physical facility may have more than one enclave present. Local and remote elements that access resources within an enclave must satisfy the policy of that enclave. A single enclave may span a number of geographically separate locations with connectivity via commercially purchased point-to-point communications (e.g., T-1, T-3, Integrated Services Digital Network (ISDN)) along with WAN connectivity such as the Internet. These concepts are represented in Figure 1-5.



**Figure 1-5 Enclave Boundaries Framework Area**

The enclave boundary is the point at which information enters or leaves the enclave or organization. Many organizations have extensive connections to networks outside their control. Therefore, a layer of protection is needed to ensure the information entering does not affect the organization's operation or resources, and that the information leaving is authorized.



Many organizations employ multiple types of external network connections through the enclave boundary. These include:

- Connections to external networks (such as the Internet) to exchange information with another enclave or to access data on a network.
- Three types of connections to remote users—dial-up access via the public telephone network, connection to an Internet Service Provider (ISP) by direct connection (cable modem) or by dial up access, and dedicated line connectivity through a Telecommunications Service Provider (TSP) (see also Figure 1-3).
- Connections to other local networks operating at different classification levels.

Each connection requires different types of solutions to satisfy both operational and IA concerns. Internets invite access through the boundary, with security only as good as the entire network through which the data is being transported.

## Networks and Infrastructures

The *network and infrastructure* of these networks provide connectivity between enclaves; they contain Operational Area Networks, (OAN), Metropolitan Area Networks (MAN), Campus Area Networks (CAN), and LANs, extending coverage from broad communities to local bases. The transport networks contain the information transmission components (e.g., satellites, microwave, other Radio Frequency (RF) spectrum, and fiber) to move information between the network nodes (e.g., routers and switches). As depicted in Figure 1-6, other important components of the network infrastructure are network management, domain name servers, and directory services.

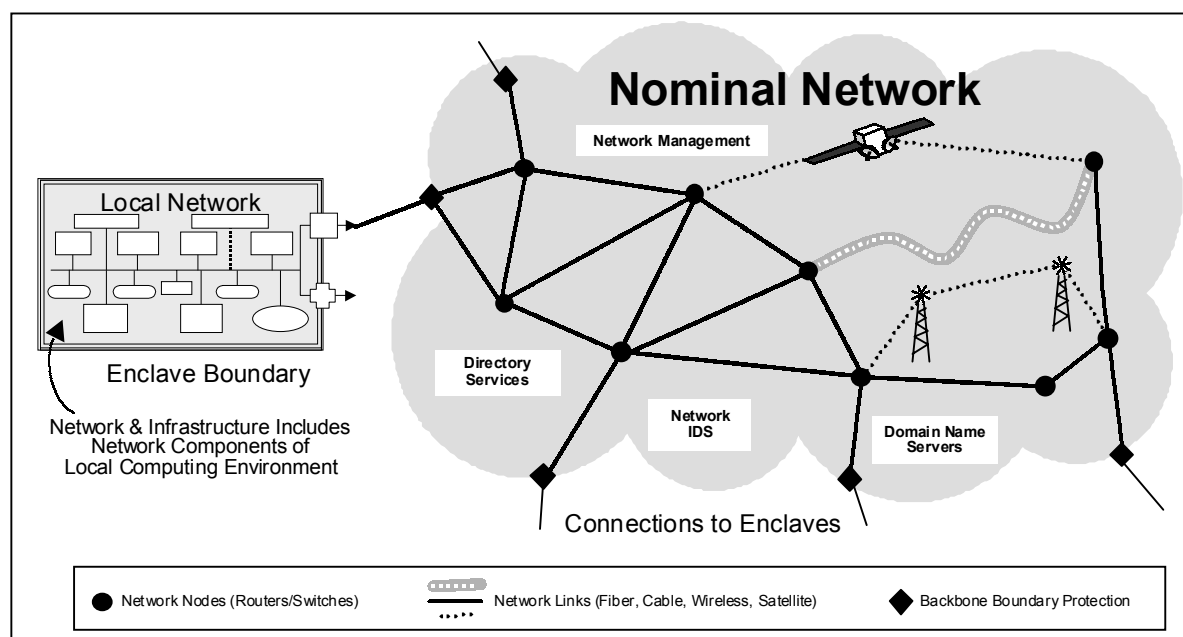


Figure 1-6 Network and Infrastructure Framework Area

The typical types of transport networks and services used by the government and industry now, and that will be used in the future, can be logically grouped into three areas:

- 1) Public/commercial networks and network technologies.
- 2) Dedicated network services.
- 3) Government owned and operated.

The public/commercial networks used by both the private sector and government include the Internet, the Public Switched Telephone Network (PSTN), and wireless networks. Wireless networks include: cellular, satellite, wireless LAN, and paging networks. Access to networks is gained typically through telecommunications service providers. These public networks are wholly owned and operated by these private sector providers.

To obtain dedicated network services, the Government has structured a number of network service contracts that procure network services. These include the Federal Wireless Service and FTS 2000. Public network providers provide access to networks through an arrangement with the Government. Private sector organizations obtain telecommunications services in a similar manner, leasing and purchasing dedicated commercial telecommunications services.

Several government organizations own and operate networks. For example, the Department of Energy's Energy Science Network (ESNet), the Federal Aviation Administration's Agency Data Telecommunications Network (ADTN), and the DoD Secret Internet Protocol Router Network (SIPRNET). These networks may start as private networks, go through leased or public networks, and terminate as private networks. They also include totally owned and operated networks such as MILSTAR. Appendix C provides additional information on this category of networks.

## Supporting Infrastructures

Also present in the information technology environment are *supporting infrastructures* that provide the foundation upon which IA mechanisms are used in the network, enclave, and computing environments for securely managing the system and providing security enabled services. Supporting infrastructures provide security services for: networks; end-user workstations; servers for web, applications, and files; and single-use infrastructure machines (e.g., higher-level Domain Name Server (DNS) services, higher-level directory servers). The two areas addressed in the IATF are key management infrastructure (KMI), which includes Public Key Infrastructures (PKI), and detect and respond infrastructures.

## Key Management Infrastructure

A KMI provides a common unified process for the secure creation, distribution, and management of the public key certificates and traditional symmetric keys that enable security services for the network, enclave, and computing environment. These services enable the identities of senders and receivers to be reliably verified, and the information to be protected from unauthorized disclosure and modification. The KMI must support controlled interoperability for users, consistent with established security policies for each user's community.

## Detect and Respond

The detect and respond infrastructure enables rapid detection of and reaction to intrusions. It also provides a “fusion” capability so one incident can be viewed in relation to others. This allows analysts to identify potential activity patterns or new developments. In most organizations that implement a detect and respond capability, local centers monitor local operations and feed a larger regional or national center. The infrastructure required includes technical solutions such as intrusion detection, and monitoring software; and a cadre of skilled specialists, often referred to as a Computer Emergency Response Team (CERT).

### 1.3.5 Nature of Cyber Threats

Information systems and networks offer attractive targets. They should be resistant to attack from the full range of threat-agents—from hackers to nation states—and they must limit damage and recover rapidly when attacks do occur.

The IATF considers five classes of attacks:

- 1) Passive.
- 2) Active.
- 3) Close-In.
- 4) Insider.
- 5) Distribution.

The key aspects of each class of attack are summarized in Table 1.1.

**Table 1-1 Classes of Attack**

Attack	Description
Passive	Passive attacks include traffic analysis, monitoring of unprotected communications, decrypting weakly encrypted traffic, and capturing authentication information (e.g., passwords). Passive intercept of network operations can give adversaries indications and warnings of impending actions. Passive attacks can result in the disclosure of information or data files to an attacker without the consent or knowledge of the user. Examples include the disclosure of personal information such as credit card numbers and medical files.
Active	Active attacks include attempts to circumvent or break protection features, introduce malicious code, or steal or modify information. These include attacks mounted against a network backbone, exploitation of information in transit, electronic penetrations into an enclave, or attacks on an authorized remote user when attempting to connect to an enclave. Active attacks can result in the disclosure or dissemination of data files, denial of service, or modification of data.
Close-in	Close-in attack is where an <i>unauthorized</i> individual is in physical close proximity to networks, systems, or facilities for the purpose of modifying, gathering, or denying access to information. Physical close in proximity is achieved through surreptitious entry, open access, or both.
Insider	Insider attacks can be malicious or non-malicious. Malicious insiders have the intent to eavesdrop, steal or damage information, use information in a fraudulent manner, or denying access of other authorized users. Non-malicious attacks typically result from carelessness, lack of knowledge, or intentionally circumventing security for non-malicious reasons such as to “get the job done.”
Distribution	Distribution attacks focus on the malicious modification of hardware or software at the factory or during distribution. These attacks can introduce malicious code into a product such as a back door to gain unauthorized access to information or a system function at a later date.

The relationship of these attack classes to the technology framework areas is shown in Figure 1-7. Subsequent sections of the IATF will provide an overview of the IA strategy for countering or mitigating the effects of these attacks.

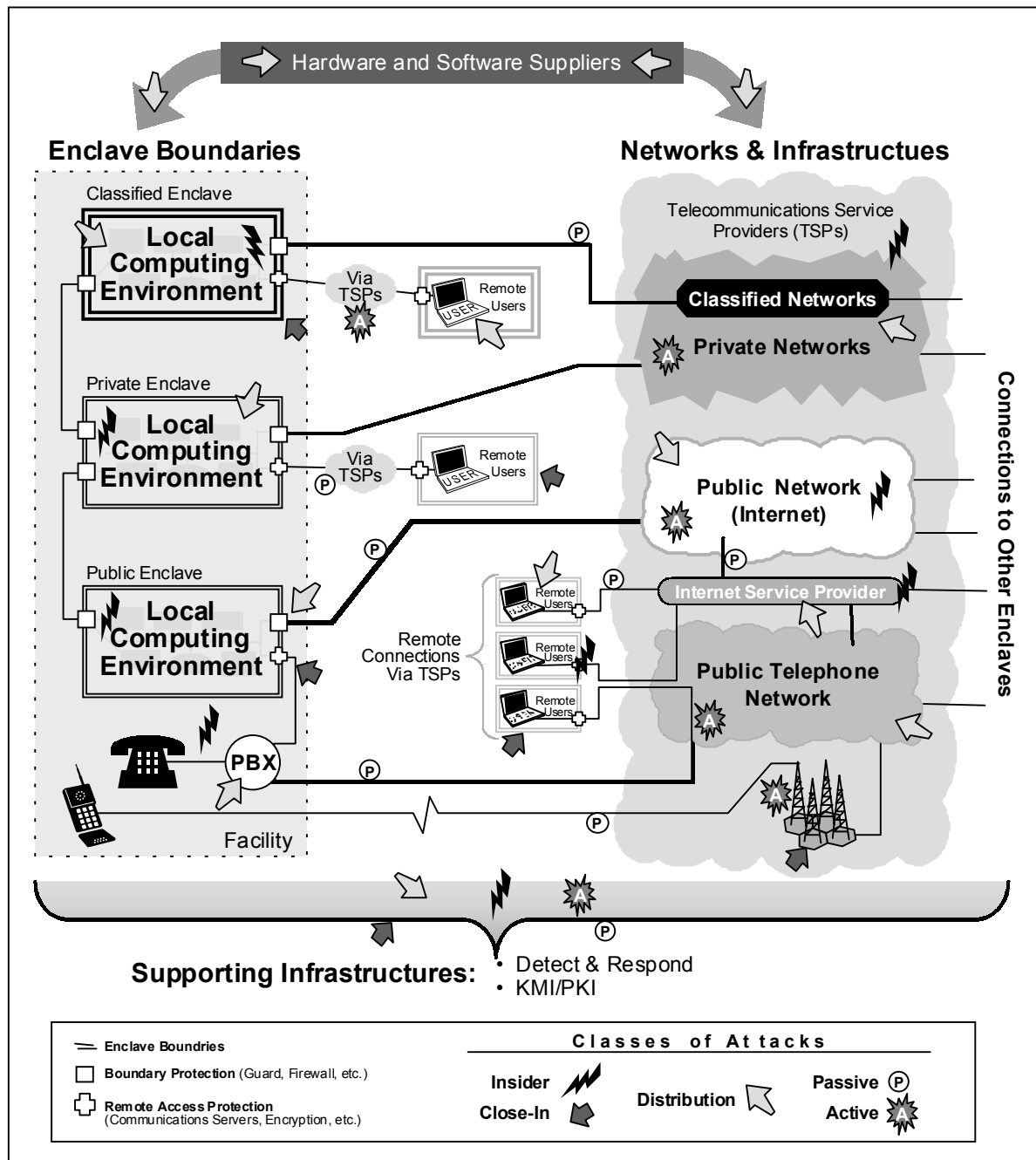


Figure 1-7 Classes of Attacks on the Information Infrastructure

## 1.4 Defense-in-Depth

The Department of Defense (DoD) has led the way in defining a strategy, called *Defense-in-Depth*, to achieve an effective IA posture. The underlying principles of this strategy are applicable to any information system or network, regardless of organization. Essentially, organizations address information assurance needs with *people* executing *operations* supported by *technology*.

Figure 1-8 illustrates the principal aspects of the *Defense-in-Depth* strategy—personnel, technology, and operations, outlined as follows.

- People
  - ♦ Training
  - ♦ Awareness
  - ♦ Physical security
  - ♦ Personnel security
  - ♦ System security administration
- Technology
  - ♦ Defense-in-Depth technology framework areas
  - ♦ Security criteria
  - ♦ IT/IA acquisition
  - ♦ Risk assessments
  - ♦ Certification and Accreditation
- Operations
  - ♦ Assessments
  - ♦ Monitoring
  - ♦ Intrusion detection
  - ♦ Warning
  - ♦ Response
  - ♦ Reconstitution

Of the three principal aspects of this strategy, the IATF focuses on technology and on providing a framework for providing overlapping layers of protection against cyber threats. By this approach, a successful attack against one layer or type of protection does not result in the compromise of the entire information infrastructure.

Other policies, procedures, and frameworks are focused on addressing the people and operations aspects of a *Defense-in-Depth* strategy.

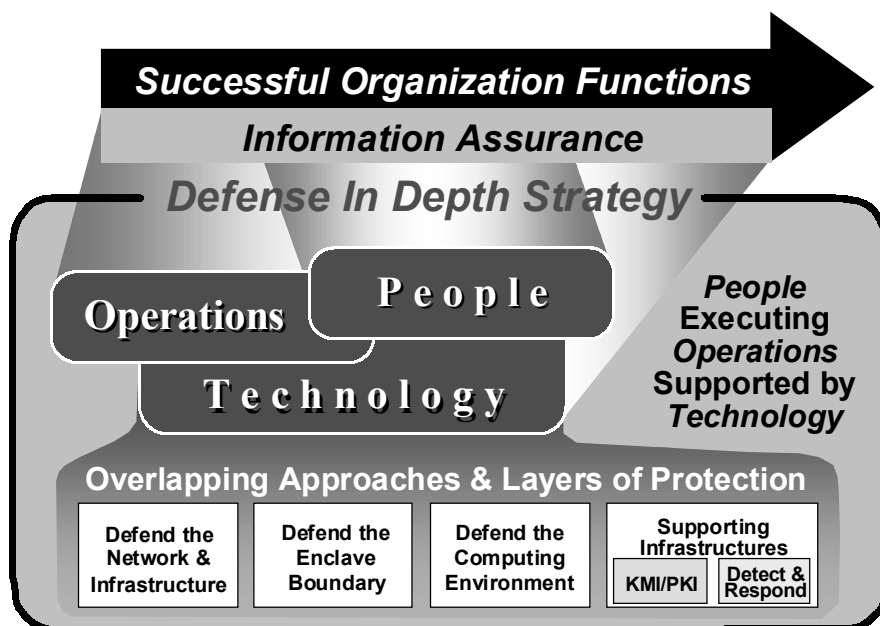


Figure 1-8 Principal Aspects of the Defense-in-Depth Strategy

### 1.4.1 Defense-in-Depth and the IATF

Information infrastructures are complicated systems with multiple points of vulnerability. To address this, the IATF has adopted the use of multiple IA technology solutions within the fundamental principle of the *Defense-in-Depth* strategy, that is, using layers of IA technology so-

lutions to establish an adequate IA posture. Thus, if one protection mechanism is successfully penetrated, others behind it offer additional protection. Adopting a strategy of layered protections does not imply that IA mechanisms are needed at every possible point in the network architecture. By implementing appropriate levels of protection in key areas, an effective set of safeguards can be tailored according to each organization's unique needs. Further, a layered strategy permits application of lower assurance solutions when appropriate, which may be lower in cost. This approach permits the judicious application of higher assurance solutions at critical areas, (e.g., network boundaries).

The *Defense-in-Depth* strategy organizes these requirements into four principle areas of focus:

- Defend the Network and Infrastructure.
- Defend the Enclave Boundary.
- Defend the Computing Environment.
- Supporting Infrastructures.

These four areas of focus for the Defense-in-Depth strategy parallel the four framework areas discussed in Section 1.3.4.

## 1.5 IATF Organization

This *framework document* has been assembled to present the technology aspects associated with the Defense-in-Depth framework areas; each of the four areas is presented in a separate chapter. Also present are chapters that address concerns that cut across the technology areas or address the information assurance needs of particular environments or technologies.

To focus on the needs of a diverse group of readers, the IATF is organized into three primary parts shown in Figure 1-9: Main Body and Technical Sections, Executive Summaries, and Protection Profiles.

The main body of the IATF (Chapters 1 through 4) provides the general IA guidance that information system users, security engineers, security architects, and others can use to gain a better understanding of the IA issues involved in protecting today's highly interconnected information systems and network backbones. The technical sections (Chapters 5 through 9 and Appendices A through E) provides specific requirements and solutions for each of the Defense-in-Depth areas. It also offers the government and private research communities a perspective on technology gaps that exist between today's best available protection solutions and the desired IA capabilities.

For users and security engineers looking for more definitive guidance, the *Executive Summaries* portion of the IATF provides outlines of the threats, requirements, and recommended solutions for a variety of specific protection needs in specific environments. The goal of this collection of Executive Summaries is to offer quick reference guides (each summary is targeted to be fewer than three pages in length) that users and security engineers can peruse to find scenarios similar or identical to their own IA challenges.

Executive Summaries are under development and will be included in a future release of the IATF. For this version of the IATF, an outline illustrating the content of an Executive Summary is provided in Appendix F. In identifying IA solutions, the Executive Summaries will point to the documentation sources (e.g., specifications and protection profiles) containing the set of testable requirements satisfying the user need.

The third part of the IATF are referenced Protection Profiles. Protection profiles capture the assurance requirements and functionality for a system or product. Protection profiles employ the international standard Common Criteria language and structure. Appendix G, a placeholder for a future section of the IATF, will contain an index of available protection profiles and may include key protection profiles as well.

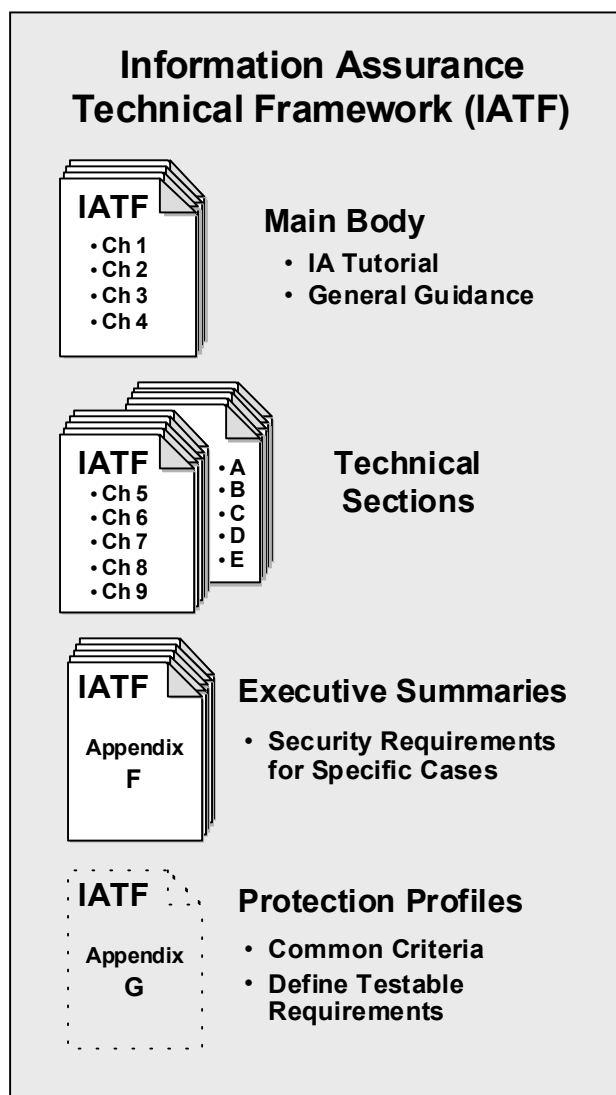


Figure 1-9 Composition of the IATF

This page intentionally left blank



# Chapter 2

## Defense-in-Depth Objectives Overview

### 2.1 Overview

The need to operate safely—*securely*—is not new. There is a continuing need for exchange of information—of varying levels; among individuals and organizations; with different clearances, missions, and needs—to support successful accomplishment of the customer's diverse missions. In parallel, the need to operate efficiently is becoming more important in the age of downsizing organizations and tightening fiscal budgets. Thus, as organizations rely increasingly on information and communications systems, as organizations strive for efficiency through shared resources, and as the people who perpetrate the threats become more numerous and more capable, the Information Assurance (IA) posture of systems and organizations grows ever more important. Deliberate investments of time, resources, *and attention* to implementing *and maintaining* an effective IA posture has never been more important or more challenging.

In implementing an effective and enduring IA capability, in adopting a *Defense-in-Depth* strategy for IA, organizations should consider:

- Effectiveness of information protection required based on the value of the information to the organization and the potential impact that loss or compromise of the information would have on the organization's mission/business. IA decisions should be risk analysis based, keyed to the organization's operational objectives.
- A composite approach, based on balancing protection capability with cost, performance, operational impact, and changes to the operation itself based on today's and tomorrow's operations and environments.
- A composite approach, drawing from all three facets of Defense-in-Depth—people, operations, and technology. Technical mitigations are of no value without trained people to use them and operational procedures to guide their application.
- A comprehensive program of education, training, practical experience, and awareness is needed. Professionalization and certification licensing provides a validated, recognized, expert cadre of system administrators.
- Exploiting available Commercial Off-The-Shelf (COTS) products, relying on in-house development for those items not otherwise available.
- Planning and following a continuous migration approach to both take advantage of evolving information processing and network capabilities—both functional and security related—and to ensure adaptability to changing organization needs and operating environments. IA protection is not static; it is a continuous adaptation.
- Periodically assess the IA posture of the information infrastructure. Technology tools, such as automated scanners for networks, can assist in vulnerability assessments.

- Not just the actions of those with hostile intent, but also the inadvertent or unwitting occurrences and the impact of natural events.
- Adherence to commonality, standardization, procedures, policies, and interoperability.
- Judicious use of emerging technologies, balancing enhanced capability with increased risk.
- Employing multiple types of mitigations, overlapping protection approaches to counter anticipated events such that loss or failure of a single barrier does not compromise the overall information infrastructure.
- Implementing *and holding* to a robust IA posture—one that can cope with the unanticipated and unexpected.
- Ensuring only trust-worthy personnel have physical access. Some methods are appropriate background investigations, security clearances, credentials, and badges.
- Monitor vulnerability listings and implement fixes, ensure security mechanisms are interoperable, keep constant watch over the security situation and mechanisms, properly employ and upgrade tools and techniques, and deal rapidly and effectively with issues.
- Incident information from intrusion detection should be reported through established procedures to authorities and specialized analysis and response centers.

The dominant need of the user community is ready access to the information and information infrastructure needed to support their operational objectives. This requires the use of robust information processing technology and reliable connectivity. IA enables these capabilities by providing organizations with the capacity to maintain adequate protection of their information.

The framework document focuses on the technology aspects of Defense-in-Depth. When developing an effective IA posture, all three components of the Defense-In-Depth strategy— people, technology, and operations need to be addressed.

For this framework document, the presentation of IA technology objectives and approaches for the information infrastructure are organized around the four Defense-in-Depth technology focus areas: Defend the Computing Environment, Defend the Enclave Boundary, Defend the Network and Infrastructure, and Supporting Infrastructures. The technology objectives and approaches in these focus areas, explained in the sections that follow, address the needs of both the private and public, civil and military sectors of our society.

## 2.1.1 Defend the Computing Environment Objectives

Users have a requirement to protect internal system applications and servers. This includes Identification and Authentication (I&A), access control, confidentiality, data integrity, and non-repudiation security services for the variety of legacy and emerging applications within system high environments. The objectives in meeting this requirement are to:

- Ensure that clients, servers, and applications are adequately defended against denial of service, unauthorized disclosure, and modification of data.

- Ensure the confidentiality and integrity of data processed by the client, server, or application whether both internal and external to the enclave.
- Defend against the unauthorized use of a client, server, or application.
- Ensure that clients and servers follow secure configuration guidelines and have all appropriate patches applied.
- Maintain configuration management of all clients and servers to track patches and system configuration changes.
- Ensure that a variety of applications can be readily integrated with no reduction in security.
- Ensure adequate defenses against subversive acts of trusted people and systems, both internal and external.

## 2.1.2 Defend the Enclave Boundary Objectives

Organizations have a requirement to protect their information infrastructures they connect to private or public networks for the purpose of obtaining information and services from those networks. This means they must protect their infrastructure, such as their local computing environment, from intrusion. A successful intrusion could result in the compromise of availability, integrity, or confidentiality. The objectives in meeting this requirement are to:

- Ensure that physical and logical enclaves are adequately protected.
- Enable dynamic throttling of services in response to changing threats.
- Ensure that systems and networks within protected enclaves maintain acceptable availability and are adequately defended against denial of service intrusions.
- Ensure that data exchanged between enclaves or via remote access is protected from improper disclosure.
- Provide boundary defenses for those systems within the enclave that cannot defend themselves due to technical or configuration problems.
- Provide a risk-managed means of selectively allowing essential information to flow across the enclave boundary.
- Provide protection against systems and data within the protected enclave being undermined by external systems or forces.
- Provide strong authentication, and thereby authenticated access control, of users sending or receiving information from outside their enclave.

## 2.1.3 Defend the Network and Infrastructure Objectives

Organizations have a requirement to protect their network and infrastructure so that information services are maintained and information—public, private, or classified—is not unintentionally disclosed or altered. The objectives in meeting this requirement are to:

- Ensure that all data exchanged over Wide Area Networks (WANs) is protected from disclosure to anyone not authorized access to the network.
- Ensure that WANs supporting mission critical and mission support data provide appropriate protection against denial of service attacks.
- Protect against the delay, misdelivery, or non-delivery of otherwise adequately protected information.
- Protect from traffic flow analysis
  - User traffic
  - Network infrastructure control information.
- Ensure protection mechanisms do not interfere with otherwise seamless operation with other authorized backbone and enclave networks.

## 2.1.4 Supporting Infrastructures Objectives

Supporting infrastructure is an enabling technology for the other Defense-in-Depth areas. This area provides the key management and detect and respond aspects of Defense-in-Depth. Such supporting infrastructure components are needed to be able to detect and respond such as intrusion detection systems, audit, configuring the system, or collecting data needed for an investigation. The objectives in meeting this requirement are to:

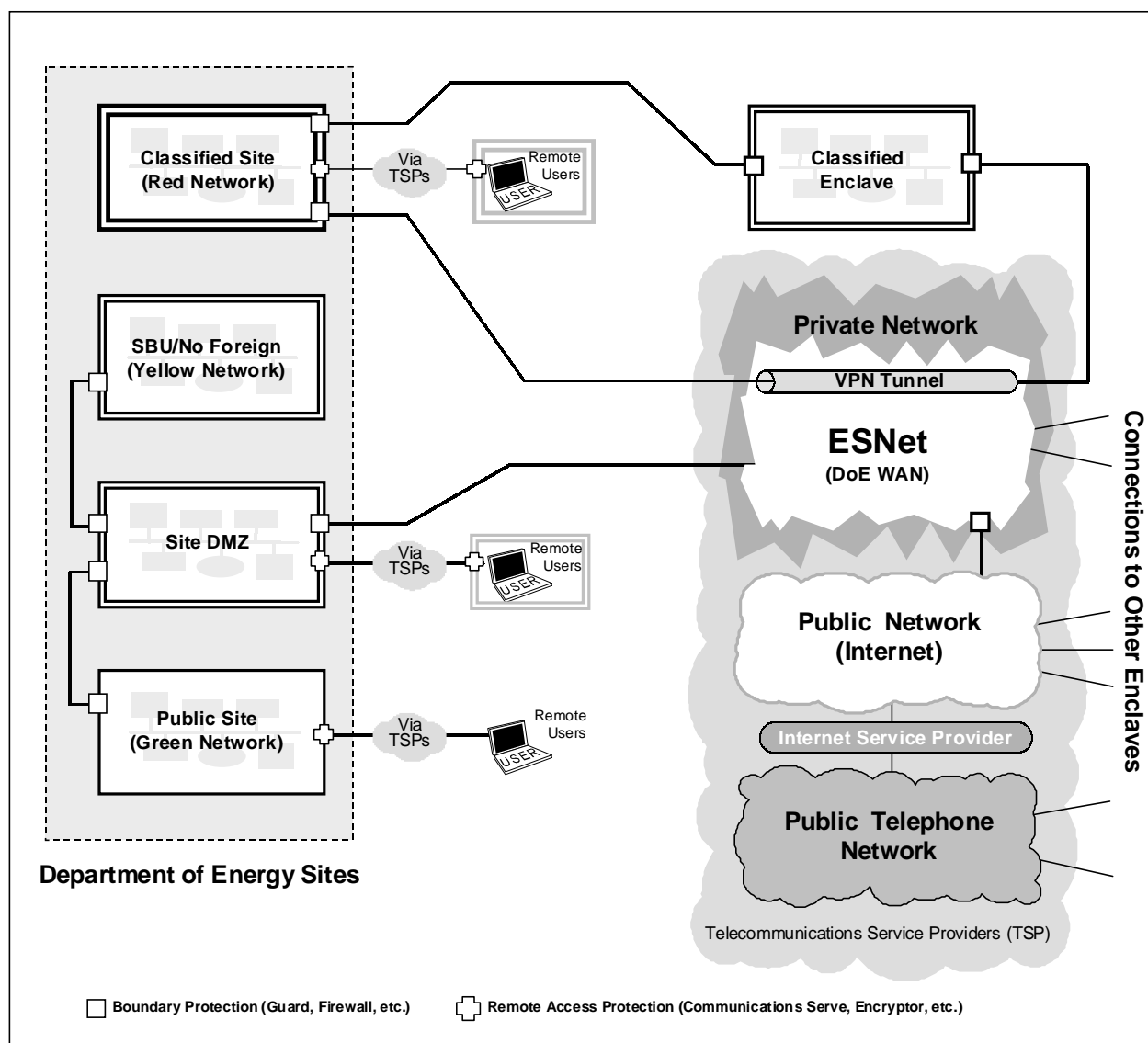
- Provide a cryptographic infrastructure that supports key, privilege, and certificate management; and that enables positive identification of individuals using network services.
- Provide an intrusion detection, reporting, analysis, assessment, and response infrastructure that enables rapid detection and response to intrusions and other anomalous events, and that enables operational situation awareness.
- Plan execution and reporting requirements for contingencies and reconstitution.

## 2.2 Examples of User Environments

The following subsections introduce example customer computing environments and depict how they may interconnect with other organizational enclaves. The IATF technologies and suggested solutions are applicable to the following computing environments.

### 2.2.1 Federal Computing Environment

The interconnection of Department of Energy (DOE) research facilities, weapons labs, regional Operations Offices, and academic facilities is an example of a Federal Computing Environment. The DOE information infrastructure is interconnected via several DOE WANs, one of which is the Energy Science Network (ESNet).



**Figure 2-1 Example Federal Computing Environment—DOE**

ESNet is a high-performance data communications backbone that provides DOE widespread support for research and mission critical applications. It supports both classified and unclassified DOE mission-oriented networking for scientists, engineers, and their administrative support. The ESNet consists of an Asynchronous Transfer Mode (ATM) backbone and multiple Local Area Networks (LANs) interconnected to establish a global network capability. ESNet permits virtual network architectures so that virtual networks can be layered on top of the existing network while running totally independent on the host network (i.e., ESNet). One example of a DOE virtual network hosted on ESNet is SecureNet, a classified DOE support network. The virtual private network, SecureNet, provides a connection between three ASCI teraflop supercomputers, DOE headquarters, and other defense program facilities across the United States (U.S.). As a result, scientists and researchers at any of these DOE sites have on demand access to the super computers.

Figure 2-1 contains a conceptual diagram of a typical DOE site and shows an extension of the broader DOE Computing Environment. The typical DOE site has two primary networks (three, if the site processes classified information).

The primary networks include a “Green” unclassified or public network, a “Yellow” or Sensitive But Unclassified/No-Foreign (SBU/NOFORN) network, and a “Red” or classified network. The Green, Yellow and Red networks may consist of one LAN, or multiple sub-networks. The typical DOE site has implemented a Demilitarized Zone (DMZ) or Information Protection Network (IPN) that acts as the single point of entry into the site and defends the enclave boundary or external connection(s). Within the yellow and red LANs, virtual networks are established to support various mission functions within the site. Physical isolation is primarily used to maintain the confidentiality and the integrity of classified data. Carefully controlled connectivity is provided between the Red network, Yellow network, and ESNet when data transfer outside the enclave is required.

All public information, web-serves, and non-sensitive information is located on the Green network which is normally protected by the site’s DMZ resources. Remote access to the site will be established via the DMZ. A typical DOE site obtains Internet access via the ESNet connection.

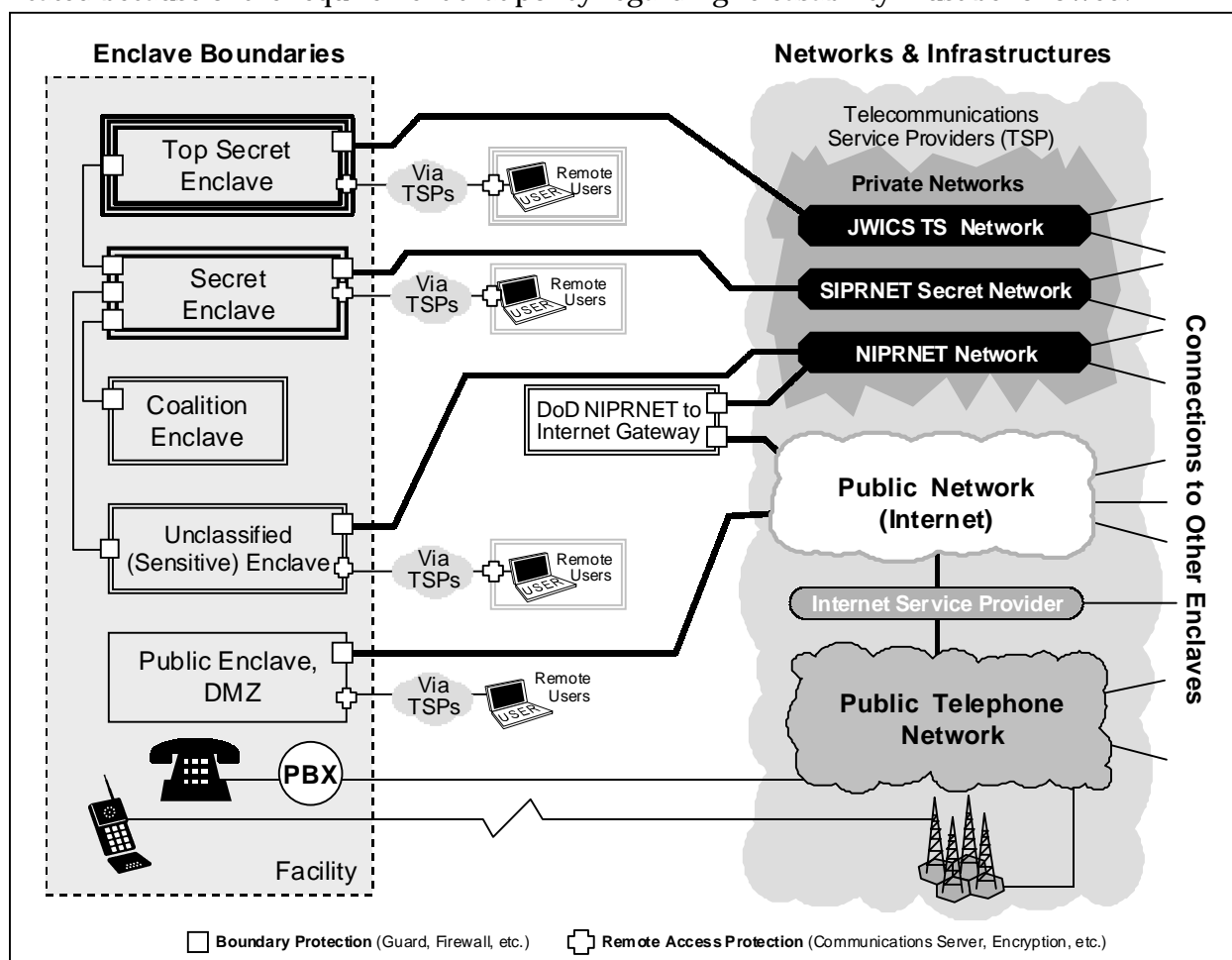
## 2.2.2 Department of Defense Computing Environment

The Defense Information Infrastructure (DII) environment is an example of one of the U.S. Government’s largest and most complex information infrastructures. The DII supports over 2 million primary users (with extensions to an additional 2 million users). Included within the DII are some 200 command centers and 16 large data centers, called, Defense Megadata Centers. The basic user environments are enclaves (physically protected facilities and compounds), incorporating over 20,000 local networks and some 4,000 connections to a backbone network. The DII also supports over 300,000 secure telephone users.

The DII implements a number of global, virtual networks that support a range of mission functions, e.g., logistics, intelligence, using WANs such as the Joint Worldwide Intelligence Communications System (JWICS) and the Secret Internet Protocol Router Network (SIPRNET) for global connectivity. In the past, this information infrastructure was built using dedicated networks and customized information systems; today, Department of Defense (DoD) is almost totally dependent on commercial services within the National Information Infrastructure (NII) and the broader global information infrastructure.

Figure 2-2 is a system context diagram of a typical user site or facility and shows an extension of the broader DII structure. The typical user facility has a number of local area networks that support the mission functional areas. Today, physical isolation is primarily used to maintain the confidentiality and the integrity of different classification levels of traffic. Within these isolated LANs, virtual networks are established to support the various mission functions within the enclave. Carefully controlled connectivity is provided between different classification levels of networks when boundaries are required.

For instance, DoD organizations have robust, worldwide intelligence systems operating at Top Secret-Sensitive Compartmented Information (TS-SCI) that carry significant levels of unclassified traffic. This supports their need to communicate with others within the intelligence community. Within the same TS-SCI enclaves, customers have Secret and Unclassified systems with less than robust connectivity to non-intelligence community users. To reach a mixed community of users, unclassified information may have to flow over separate Unclassified, Secret, and TS-SCI systems. Moving information between these systems (enclaves) is complicated because of the requirement that policy regarding releasability must be followed.



**Figure 2-2 Example Federal Computing Environment—DoD**

**This page intentionally left blank**



# Chapter 3

## Information Systems Security Engineering Process

Information Systems Security Engineering (ISSE), as described in this chapter, is the process for addressing a user's information protection needs and is part of systems engineering, systems acquisition, risk management, certification and accreditation, and life-cycle support processes. The process provides the basis for the background information, technology assessments, and guidance contained in the remainder of the Information Assurance Technical Framework (IATF) document and ensures that security solutions are effective and efficient.

### 3.1 Introduction to Chapter 3

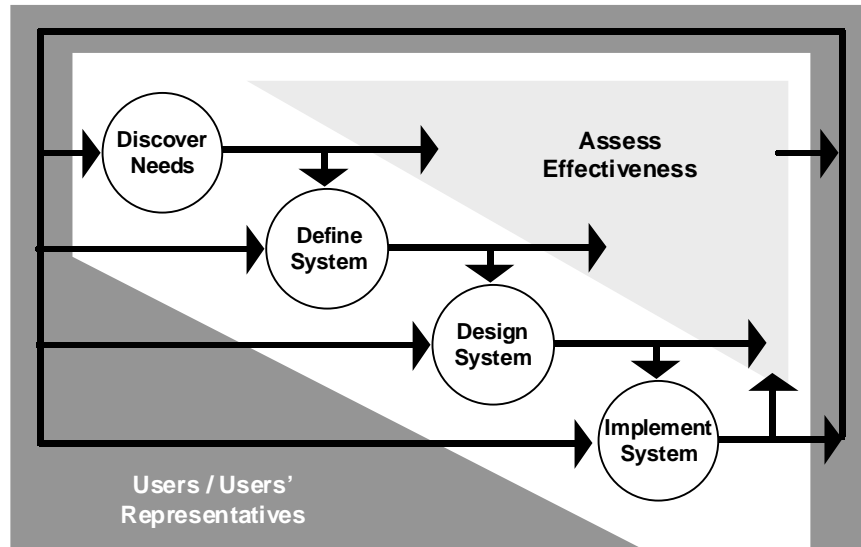
The purpose of Chapter 3 is to introduce the ISSE process and to describe the relationship between ISSE and other processes related to information systems. The ISSE process provides the background for tailoring and applying the solutions found in subsequent IATF chapters.

To achieve this purpose, Chapter 3 is organized into three sections. After the Introduction, Section 3.2 (Information Systems Security Engineering) describes the systems engineering and ISSE processes. Both share common elements—discovering needs, defining system functionality, designing system elements, producing and installing the system, and assessing the effectiveness of the system. Section 3.3 (Relation of ISSE Process to Other Processes) describes the relationship of ISSE to other systems processes—systems acquisition, risk management, certification and accreditation, and life-cycle support processes. Section 3.3 also provides suggestions on how the Common Criteria might be used to support the ISSE process.

### 3.2 Information Systems Security Engineering

#### 3.2.1 Purpose and Scope of Section 3.2

Section 3.2 (Information Systems Security Engineering) of Chapter 3 summarizes both the systems engineering and the ISSE processes. Figure 3-1 shows the major activities of the systems engineering process. The figure also reflects the relationships between activities in the process. Arrows indicate the flow of information between the activities but not their sequence or timing. Through *Assess Effectiveness*, the work products of each activity are evaluated to ensure that the system will meet the users' needs by performing the required functions to the required quality standard in the intended environment.



**Figure 3-1 Systems Engineering Process**

### 3.2.2 Overview of the Systems Engineering Process as a Basis for ISSE

The systems engineering activities described in Figure 3-1 flow in the following general manner:

- Discover Needs of mission or business,
- Define System Functionality,
- Design the System,
- Implement the System, and
- Assess Effectiveness.

The systems engineering process exercises the following principle—separate the “problem space” from the “solution space.” The problem space represents the constraints, risks, policies, and other limits on the concept of the solution. The solution space represents the activities completed and products created while developing the system to meet users needs. The progress of the systems engineering activities and work products, represented by the solution space, toward the defined and agreed-to goal is constantly evaluated for effectiveness and violations of problem space conditions. These evaluations are the basis for making the necessary corrections to the problem and solution spaces. Separating the problem space from the solution space allows effective solutions that are in agreement with physical laws and man-made policies to be conceptualized and created.

### 3.2.2.1 Discover Needs

The systems engineering process begins with the determination of the user's mission needs, relevant policies, regulations, and standards in the user environment. The system engineer should identify all of the users; the nature of their interaction with the system; and their roles, responsibilities, and authorities in each stage of the system life-cycle. The needs should come from the user's perspective and should not overly constrain the design or implementation of the system. An essential element of this process is obtaining documentation that captures in the user's language a description of the mission or desired capability, the current capability deficiency or market opportunity, the (market) environment, and a description of how the system will be used to achieve mission objectives or market position.

The major inputs, mission description and policies, for defining the mission needs are described in the following subsections. The organization responsible for performing the mission generates the mission description, but a higher level organization gives the directions for the derived mission. Finally, policy acts as a constraint, typically affecting many areas in the system life-cycle from system definition, design, and implementation to operation, support, and disposal.

These factors are used to develop a Mission Needs Statement (MNS) and a Concept of Operations (CONOPS) that will drive the definition, design, and implementation of the features for the system to operate successfully within the context of the system environment.

#### 3.2.2.1.1 Mission/Business Description

*Merriam-Webster's Collegiate Dictionary* (<http://www.m-w.com/>) [1] defines mission as: "4a - a specific task with which a person or a group is charged, b (1): a definite military, naval, or aerospace task ... ."

The mission and functions of a specific organization must be considered with respect to the mission and functions of its parent organization as well as those of other organizations that it impacts. The organization's important assets (e.g., the categories of information and/or available resources to include in-place network resources, such as information processing/storage resources) must be addressed when explaining why a system is needed.

In describing the mission, it is important to describe the mission's environment—the circumstances under which the mission will be accomplished. The mission environment may be competitive, as in the case of business, or hostile, as in the case of active military engagement. The mission environment may be complex—describing the circumstances, conditions, and engagements routinely expected as well as exceptional situations and rules for, or indicators of, change from the routine to the exceptional environment.

The established roles and responsibilities of stakeholders in the mission must also be addressed. An example would be a U.S. military organization whose specific function is to test certain equipment used to provide defense from armed attacks by a foreign organization. How the information is managed during and after the tests needs to be described. Another example is that of a financial institution. One aspect of a financial institution's operation is to ensure that funds are correctly transferred from one account to another. The institution should indicate the range of values of the individual fund transfers as well as the aggregate of all funds transferred. Each organization must describe its information management needs. Relevant questions that

must be answered include, but are not limited to, what rules govern information handling, who has access to information and information assets, and what contribution do the information and information assets make to successful mission completion. In each example, it is important for the individual organization to characterize the roles and responsibilities of individuals who use or interface with the system during the various information processing steps.

### 3.2.2.1.2 Policy Considerations

An organization must consider all the existing policies, regulations, and standards that are binding on that organization. For example, national, executive level, Department of Defense (DoD), and Navy policies may bind an U.S. Navy base. These all must be considered as inputs to the formulation of a local security or information protection policy for a particular base.

Some examples of existing policies include the following.

- DODD 8000.1, 10/27/1992, “Defense Information Management (IM) Program.” This policy directs that accurate and consistent information be made available expeditiously to DoD decision-makers to effectively execute DoD missions and to plan, acquire, develop, and implement information systems from a DoD-wide perspective to ensure consistency of information processes in and across functional areas.
- DODD 5000.1, 3/15/1996, DEFENSE ACQUISITION (administrative re-issuance incorporating Change 1, 21 May 1999). This policy directs that DoD acquisition executives [2] translate operational needs into stable, affordable programs, [3] acquire quality products, and [4] organize for efficiency and effectiveness so as to acquire quality products that satisfy the needs of the operational users with measurable improvements to mission accomplishment, in a timely manner and at a fair and reasonable price.
- DODD 4710.1, 6/21/1984, “Archaeological and Historic Resources Management.” This policy integrates the archaeological and historic preservation requirements of existing laws with the planning and management of DoD-controlled activities; seeks to minimize expenses through judicious application of options available in complying with applicable laws; and encourages practical, economically feasible rehabilitation and adaptive use of historical resources.
- DODD 4630.5, 11/12/1992, “Compatibility, Interoperability, and Integration of Command, Control, Communications, and Intelligence (C3I) Systems.” This policy directs that C3I systems be compatible, interoperable, and integrated throughout the spectrum of conflict (global long-term goal), that all C3I system be considered for joint use, and that interoperability and integration requirements be determined during requirements validation and be updated as necessary throughout the system life-cycle.

System engineers will examine the policies for relevant needs, requirements, design constraints, and other statements not expressed elsewhere that will affect the system design.

## 3.2.2.2 Define System Functionality

### 3.2.2.2.1 Objectives

During the Define System Functionality activity of system development, the system engineer must define what the system is going to do, how well the system must perform its functions, and what the external interfaces for the system are. The system engineer will also translate the natural language describing the environment in which the system will be used into engineering diagrams that define interfaces and system boundaries.

Each translation, from needs to objectives, objectives to requirements, and requirements to functions takes on language that is more engineering-like. The system engineer will generate specific objective statements in the form of *infinitive verb + object + qualifying statement* (e.g. to protect information confidentiality) that satisfy the needs by describing the expected operational results of the system. The system engineer must be able to explain logically and trace each objective to a stated need. Each objective will get a Measure of Effectiveness (MoE) that describes the conditions when the objective is satisfied. Therefore, objective statements should be unambiguous, measurable, and verifiable. When all the objectives that trace to a need are satisfied, then, if the translation from needs to objectives is done correctly and completely, the needs are satisfied.

### 3.2.2.2.2 System Context/Environment

The technical system context identifies the functions and interfaces of the system that interact with elements outside of the system boundaries. The context should address physical and logical boundaries and the general nature of the inputs and the outputs to the system. Included in the context is a description of the bi-directional flow of information, signals, energy, and material between the system and the environment or other systems. The context should indicate the type of information processing required to support the users' operational missions (e.g., peer-to-peer communication, broadcast communication, storage of information, general access, restricted access, etc.).

### 3.2.2.2.3 Requirements

Functional requirements are derived from and traced to their parent objectives. Functional requirements describe what tasks, actions, or activities the system needs to accomplish. Objective MoEs when translated to performance requirements define how well the functional requirements must be performed. In addition to stating the functional, performance, interface, interoperability, derived, and (possibly) design requirements of the system, the system engineer will have to determine with the customer the assurance requirements of the system development or upgrade. Assurance requirements affect the method of design and documentation and give the customer confidence that the system does only what its developers claim it does and nothing else. Assurances may be performance requirements and specify a strength of function or process requirements and specify a method for verifying or validating a system's correctness (of design or fitness for purpose, respectively). The system engineer will likely have many trade-offs to consider between allocating performance to one set of requirements or another to determine a set of performance requirements that will satisfy the needs and represent

an acceptable level of risks, life-cycle costs, and schedule requirements. Performance requirements typically take the form of stating:

- Quality - How well?
- Quantity - How many, How much cost per system?
- Coverage - How far, How much covered?
- Timeliness - How frequent, How responsive?
- Readiness - Reliability, Maintainability, Availability, Producibility.

Internal interface, external interface, and interoperability requirements are important requirements that may be derived from the concept of the system's components' interactions with each other or with other systems and the environment. Furthermore, policy may dictate certain interface, interoperability, and design requirements. The system engineer may have to derive further requirements that are necessary for system functioning.

When all of the requirements are captured, the system engineer will review them with other system stakeholders for correctness, completeness, coherence, interdependence, conflicts, and testability. Correct requirements that properly translate the objectives should not be extreme or vague. An extreme requirement is a dubious requirement that calls for some system property that is so strong that it modifies other, reasonable requirements. Each requirement should be necessary, and the set of requirements should be sufficient to satisfy the needs of the user. Each requirement should be coherent — understood by the users, customers, and developers to mean the same specific thing. The system engineer should resolve conflicts between requirements by negotiating with the other system stakeholders which requirements to eliminate or modify. Except where policy requires a certain design, requirements should be independent of implementation. The user should prioritize the set of requirements. Lower priority “desirable” requirements can be traded for time, money, reduced risk, or reduced scope if necessary when conflicts occur. It is important to have stakeholder agreement on the set of requirements to be met and the nature of the requirements because the requirements are the basis for system validation and acceptance.

A Requirements Traceability Matrix (RTM) to trace needs to requirements should be created during this activity. Also, because the requirements are the basis for validation, the test plans that will exercise each requirement or set of requirements should be referenced or included in the RTM.

### 3.2.2.2.4 Functional Analysis

Functions are derived from requirements. Each requirement will generate one or more functions in the form *verb + object*. Depending on how the requirements and functions are written, the initial function derived from the requirement may be a higher order or lower order function. The main activity in this analysis is to understand how the functions relate to each other and to their environment.

There are many ways to draw a diagram describing how functions relate to each other. The simplest diagram is a text list of the functions that shows their hierarchy using indentation, numbering, and font conventions. The function list will name the function and describe what it is, what it acts on, when it is invoked, what it accepts as input, and what it outputs. For exam-

ple, the function called “Transform Energy” (in a generator, for example) will detail the input conditions and energy (gears turning to make mechanical energy input, for example), and the resulting output conditions and energy (turning gears make alternating current electrical energy output, for example). Systems engineers developing the simplest systems may find that a function list is enough, but more commonly, the function list is just the beginning of the functional analysis.

The hierarchical nature of the function list can also be diagrammed into a functional tree. The thought process for both analyses forces the system engineer to consider the relative level or the functions in relation to the relative level of system aggregation. Grouping higher level functions and their descendants maintains a high level of independence from other groups of functions. This is part of defining a modular architecture that is cohesive (each module or subsystem produces a significant system function composed of closely related lower level functions) and loosely coupled (each module or subsystem is largely independent of all the others). The system engineer must carefully consider the trade-off between cohesiveness and coupling; modular systems are easier to specify, design, build, test, replace, and upgrade as (almost) independent subsystems. This trade-off space leads to many possible system architectures. This begins the visualization of subsystems and lower level components to which functions may be allocated.

Interface descriptions may be visualized by drawing N2, functional flow block, and context diagrams. N2 diagrams show the inputs to a function on one axis (typically vertical) and the outputs of the function on another (horizontal) axis. The diagram shows how each function depends on every other function. More simple designs will show fewer inputs intersecting outputs away from the diagonal. Interfaces can also be shown in functional flow block diagrams. Functional flow block diagrams indicate complexity and process flow. Context diagrams show the system in the context of other systems that it interacts with and in the context of the environment. As lower level context diagrams are built, the specific nature of the external functional interfaces becomes clear.

### 3.2.2.3 Design System

This activity requires a multi-disciplined team to build the architecture and specify a design solution. The system engineer categorizes the solution architecture and identifies any reusable solutions that are similar. At this point, the engineer forms a team to develop the specific solution. The team selects products to implement into the solution and designs the specific solution architectures either by tailoring reusable worked solutions or by designing new solutions.

It is important to realize that the system will depend on all of its components to accomplish the mission. Therefore, optimizing one component beyond the ability of the system to use the improvement is a waste of effort and resources, but letting component performance fall too far may damage overall system health. The system design will implement the set of requirements, which includes the functional, performance, interface, interoperability, and design requirements. The work products of this activity will answer how the system will meet the customers' needs.

### 3.2.2.3.1 Functional Allocation

As the system engineer proceeds through this process, he or she must visualize what physical form the components implementing the functions will take. Some functions may be allocated to hardware, software, firmware, or people. People performing system functions will generally have defined processes, written procedures, and specified hardware, software, and firmware tools available, especially if they are to perform the function consistently well. Therefore, some functions may be shared between things and people. As functions are allocated to components, the components become responsible for satisfying the corresponding functional and performance requirements as well as a subset of the overall system constraints in the problem space. The system engineer will have to trade various architectural concepts, where functions and requirements are allocated to components, and negotiate an agreement with system stakeholders that is both conceptually and physically feasible.

At this point, the system engineer should begin planning the system verification, integration, and validation testing required to verify the design, integrate and produce a working system, and validate the requirements. Validation, verification, and integration test plans with expected results are written and linked to the requirements and architecture. It is also time for the system engineer to allocate capital, personnel, facilities, and time resources for the design, testing, logistics, and life-cycle support for the system. Most systems require formal Configuration Management (CM); CM should be imposed on the architecture. There are other systems engineering and project management activities that are required to successfully bring a complicated system into existence, but explaining them is beyond the scope of this work.

### 3.2.2.3.2 Preliminary Design

The entry conditions to this activity of system development are, at a minimum, stable agreement on system requirements and a stable architecture under CM. Once the architecture is defined and baselined, system and design engineers will have to generate specifications that detail what is to be built. The specifications must be relevant (to the needs specified through the RTM), complete, and consistent. Specification levels of detail flow from the system level down to the component level. Production and review of the higher level specifications occurs before the Preliminary Design Review (PDR). The PDR results in the high-level specifications being reviewed for completeness, conflicts, compatibility (with interfacing systems) verifiability, security risks, integration risks, and traceability (and satisfaction of) requirements. The result of the Preliminary Design activity is the Allocated System Baseline Configuration.

### 3.2.2.3.3 Detailed Design

Detailed design results in lower level product specifications, detailed engineering and interface control drawings, prototypes, detailed test plans and procedures, and detailed Integrated Logistics Support Plans (ILSPs). Specialty engineering practices, reliability, maintainability, availability, quality, safety, and producibility provide expertise and details in selecting what is bought or built and how. This activity will conclude with the system Critical Design Review (CDR)—a review of all the detailed specifications of the Configuration Items (CI) for completeness, conflicts, compatibility (with interfacing systems) verifiability, integration risks, and traceability to (and satisfaction of) requirements.



### 3.2.2.4 Implement System

The objective of the Implementation activity is to procure or produce and integrate all components for the designed system. Once this completed, the next process in this activity is to verify and validate the system for compliance with its requirements. Some very low-level design activities (e.g., design of small software routines) will usually be conducted as part of the system “build” process during this test activity. This effort also includes reviewing the readiness of the system for manufacturing.

This activity concludes with a System Verification Review (SVR) where evidence is presented that the system, as built, complies with the system requirements and satisfies the mission capability needs. Issues across all of the systems engineering primary functions must be considered and any interdependencies or trade-offs resolved.

#### 3.2.2.4.1 Procurement

An important set of decisions to be made in this activity is whether to make or buy components that meet the detailed specifications in the Design System activity. Selecting and acquiring products for integration into the solution is an activity that is based on the selected detailed design. These products can be purchased, leased, or borrowed. The decision will be based on many factors that are known (cost of the component, availability, form, fit, and function, etc.) and factors that are unknown (reliability in the particular system, risk to system performance if component performance is marginal, future availability of the component or substitutes, etc.). To document and make formal the decisions to make or buy, the system and design engineers will have to conduct trade-off studies.

#### 3.2.2.4.2 Build

During this activity, the developed system approach is translated into a stable, producible, and cost-effective system design. For information systems, this translation includes any necessary production-level software, hardware, or firmware.

Once procurement and delivery of system components has occurred, the next activity in the Systems Engineering (SE) process is to assemble or build the system. Before this takes place, the system components should be verified as corresponding to the specifications in the design activity. Once verified, system assembly can occur. The system assembly process should be performed in accordance with manufacturers' specifications to avoid unnecessary hindrances.

The completion of the build process will significantly impact the remaining activities. If the system is assembled correctly, the remaining processes will more accurately reflect the validity of the system's design and engineering activities. However, if system assembly is faulty, the system will not function as designed and will fail to meet the design and mission objectives.

#### 3.2.2.4.3 Test

As components are developed, they will have to be tested. The systems and design engineers will write test procedures with expected results as the design solution becomes better defined. The design engineer will perform unit testing. Verification of the design and interfaces ensures

that the produced component operates correctly. All the interfaces must be exercised fully during verification and integration testing.

Integration testing verifies higher aggregate levels of system performance. The people, tools, facilities, and capital resources required to perform system testing must be specified, budgeted, and, to the largest extent possible, built, and verified during earlier activities. Selected, acquired, or built products are integrated into the proposed solution and tested to the point where higher levels of functionality are achieved. Integration testing may result in product changes that are fed back for system redesign consideration. A system functional test report is generated documenting the positive and negative results of the testing. Integration is an activity that results with providing the customer with a system that has been fully integrated and tested to determine that the system design is verified.

Often, mission needs call for a system to be developed that is unique or that is to be placed in a poorly understood or difficult to model environment. In this case, common to the Government, validation testing must occur on the installed system unless the acceptance clauses in the contract allow for validation of the requirements in a laboratory setting. Where many identical systems are going to be deployed in well-understood and well-modeled environments, it is prudent to conduct validation testing before production and deployment. Validation testing and acceptance testing are not necessarily the same, but customer acceptance and repeat business will be based on satisfaction of the requirements.

As the system is verified, integrated, and validated, it is important to document installation, operation, maintenance, and support procedures. These procedures will be based on the requirements, architecture, design, and test results of the system “as-built” configuration. As the installation proceeds, it is important to document defects in the procedures and note how changes may impact integration and validation testing and operational procedure. The impact of installation changes on the residual risk from operating, supporting, and maintaining the system should also be assessed.

### 3.2.2.5 Assess Effectiveness

In assessing the effectiveness of the system, two major factors must be examined. First, does the system meet the needs of the mission? Second, does the system operate in the desired manner of the mission organization? There may exist preconceived expectations as to the functionality and operation of the system that must not be overlooked. Functional and operational requirements of the system and are the major aspects by which the system will be accepted or rejected. It should be noted, however, that in addition to these factors, there are a number of other elements that can affect the outcome of this evaluation including:

- Interoperability. Does the system share information correctly across external interfaces?
- Availability. Is the system available to users to enhance mission success?
- Training. What degree of instruction is required for users to be qualified in operating and maintaining the system?
- Human/Machine Interface. Does the human/machine interface contribute to users making mistakes that compromise the system and mission success?
- Cost. Is it financially feasible to construct, upgrade, and/or maintain the system?

### 3.2.3 The ISSE Process

The ISSE process is a sub-process of the system engineering process described above that focuses on information protection needs and ideally occurs in parallel with the systems engineering processes. ISSE supports the evolution, verification, and validation of an integrated and life-cycle balanced set of system product and process solutions that satisfy customer information protection needs. The ISSE process also focuses on identifying, understanding, containing, and optimizing information protection risks. ISSE activities are directed toward:

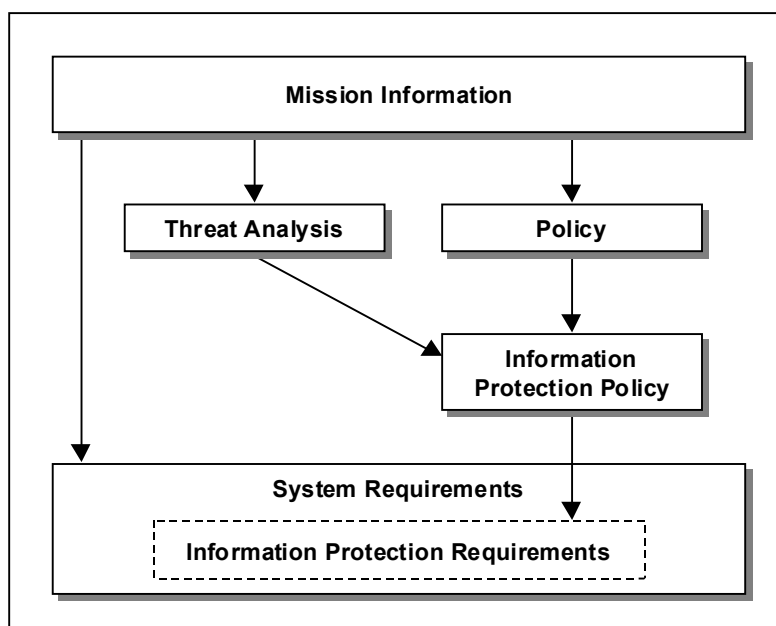
- Describing information protection needs;
- Generating information protection requirements based on needs early in the systems engineering process;
- Satisfying the requirements at an acceptable level of information protection risk;
- Building a functional information protection architecture based on requirements;
- Allocating information protection functions to a physical and logical architecture;
- Designing the system to implement the information protection architecture;
- Balancing information protection risk management and other ISSE considerations within the overall system context of cost, schedule, and operational suitability and effectiveness;
- Participating in trade-off studies with other information protection and system engineering disciplines;
- Integrating the ISSE process with the systems engineering and acquisition processes;
- Testing the system to verify information protection design and validate information protection requirements; and
- Supporting the customers after deployment and tailoring the overall process to their needs.

ISSE activities should begin with the system engineering activities to ensure that information protection is built into the overall system. Considering information protection objectives, requirements, functions, architecture, design, testing, and implementation simultaneously with the corresponding system engineering analogues allows information protection to be optimized based on the technical and non-technical considerations of the particular system.

#### 3.2.3.1 Discover Information Protection Needs

The ISSE process begins with a review of the user's mission needs, relevant policies, regulations, standards, and threats with respect to information in the user environment that was defined by the system engineers. The ISSE then identifies the users of the information systems and information, the nature of their interaction with the information systems and information, and their roles, responsibilities, and authorities in each stage of the information protection system life-cycle. The information protection needs should come from the user's perspective and not overly constrain the design or implementation of the system.

In the Information Protection Policy and the Security CONOPS, the ISSE should describe, in the user's language, how information protection supports successfully achieving the mission or desired market capability in the overall system environment. When the information protection needs of the system are discovered and described during this activity, then the information protection system will develop as an integral part of the overall system development process. Figure 3-2 shows how the analysis is done to describe how the mission, threats, and policies of the system impact the information protection requirements.



**Figure 3-2 Impact of Mission, Threats, and Policies on Information Protection Requirements**

### 3.2.3.1.1 Mission Information Protection Needs

The role of information and information systems in the larger mission and functions of a specific organization must be considered. The ISSE must consider the impact to the mission of organizational elements—people and systems—losing the use of the information systems or information that they depend on, specifically, the loss of confidentiality, integrity, availability, non-repudiation, or any combination thereof. At this point, the ISSE has begun to elicit information protection needs from the user.

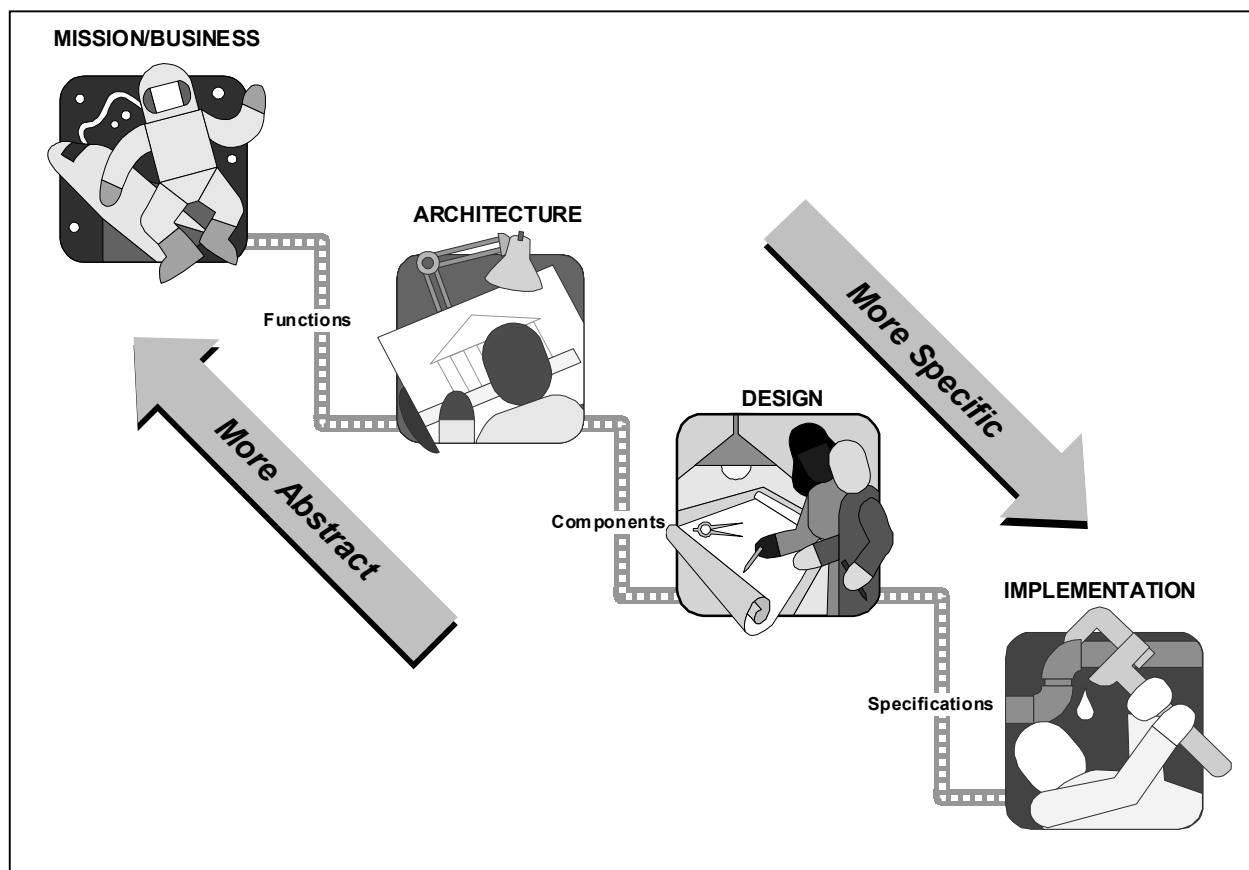
Users know best the importance of their information but usually need help in discovering their protection needs and priorities. Discovering the customer needs leads to the information protection needs in terms of what information could be used to harm the mission if it were disclosed, modified, or lost. The ISSE should be able to:

- Assist customers in modeling their information management process,
- Assist customers in defining information threats,
- Assist customers in prioritizing protection needs,
- Prepare information protection policies, and
- Achieve customer agreement.

Identifying needs is a customer interface activity performed by the ISSE to ensure that the mission/business needs include information protection needs and that the system functionality includes the information protection functionality. The ISSE brings together security disciplines, technology, and mechanisms and applies them to satisfy the protection needs of the customer.

The result is an information system that includes the information protection architecture and mechanisms that best meet the protection needs within the cost, performance, and schedule allowed by the customer.

Figure 3-3 shows the layered requirements “hierarchy” asserting that a higher layer imposes requirements on the next lower layer. The example of a good requirement depends on its position in the hierarchy. The nature of the requirements are “more specific” as you move downward in the hierarchy and “more abstract” as you move upward.



**Figure 3-3 Layered Requirements “Hierarchy”**

The ISSE must adhere to the customers' priorities in designing protection for information systems and the information that the systems perform functions on based on an assessment of the information and systems' value to the mission. The role of information and information systems in supporting the mission should be described in terms of:

- What kind of information records are being viewed, updated, deleted, initiated, or processed (classified, financial, proprietary, personal private, etc.)?
- Who or what is authorized to view, update, delete, initiate, or process information records?
- How do authorized users use the information to perform their duties?

- What tools (paper, hardware, software, firmware, and procedures) are authorized users using to perform their duties?
- How important is it to know with certainty that a particular individual sent or received a message or file?

The ISSE and the system users will have to work together on the nature of the role of information systems in furthering the users' mission. An ISSE making these decisions without user input is not likely to satisfy the users' needs.

### 3.2.3.1.2 Threats to Information Management

In terms of ISSE, the technical system context identifies the functions and interfaces of the information system that interacts with elements outside of the system boundaries. The context should address physical and logical boundaries and the general nature of the inputs and the outputs to the information system. Included in the context is a description of the bi-directional flow of the information carried on signals, energy, and material between the system and the environment or other systems. Both intended and unintended interfaces with the environment and other systems must be considered. Part of describing unintended interfaces is describing the threat environment to information and information systems. A threat is defined as the potential for circumstances in which some agent might take some action, that could cause some event, having a consequence, that could result in a harmful impact. The threat context will be described in terms of:

- Types of information,
- Legitimate users and uses of information,
- Threat agent considerations,
  - Capability,
  - Intent,
  - Willingness,
  - Motivation,
  - Damage to mission.

The context description defines some of the problem space for the information protection solution, but it also provides the underpinnings for the Risk Management process, which will be discussed in Section 3.3.3. Threats are treated more fully in Chapter 4 (Technical Security Countermeasures).

### 3.2.3.1.3 Information Protection Policy Considerations

An organization must consider all the existing information protection policies, regulations, and standards that are binding on that organization and develop a system information protection policy. The multiple layers of policies have already been described.

The most important issues an information protection policy must define are:

- Why protection is needed,
- What protection is needed, and

- Not how protection is achieved.

Just as in the systems engineering process, an organization must consider all the existing policies, regulations, and standards that are binding on that organization. For example, national, executive level, DoD, and Navy policies may bind an U.S. Navy base. These all must be considered as inputs to the formulation of a local information protection policy for a particular base.

Some examples of existing policies include the following.

- DoD Directive, 5200.28, "Security Requirements for Automated Information Systems (AISs)." It articulates the minimum security requirements for AISs as accountability, access, security training and awareness, physical controls, classification/sensitivity markings, "need-to-know" restrictions, proper handling of data throughout its life-cycle, contingency planning, risk management, and accreditation.
- Office of Management and Budget Circular A-130, Appendix III, "Security of Federal Automated Resources," and Public Law 100-235. They delineate requirements to protect all U.S. Government information systems to the level commensurate with the risk, to define roles and responsibilities of individuals authorized to have the information, and to develop and implement appropriate security plans that address continual administrative support throughout the system life-cycle.
- Executive Order 12968, "Access to Classified Information." It delineates personnel security requirements for allow individuals access to classified information.

The most important issues an organizational security policy must define are:

- The resources/assets the organization has determined are critical or need protection.
- The roles and responsibilities of individuals that will need to interface with those assets (as part of their operational mission needs definition).
- The appropriate way (authorizations) authorized individuals may use those assets (security requirements).

A multi-disciplined team of systems engineers, ISSEs, users' representatives, accreditation authorities, certifying authorities, and design specialists is needed to develop an effective organizational information protection policy. The team needs to work together to ensure that the various inputs to the policy are correctly and completely articulated, and that the resultant policy is correctly stated and consistent.

Senior management must issue the organizational information protection policy. It needs to be decisive and set a direction to enable lower level decisions to be made. The policy must be available to, and easily understood by, the entire workforce. There must be a procedure to ensure the policy is enforced throughout the organization, and the workforce must understand the organizational and personnel consequences if the policy is not enforced. Although the organizational information protection policy must be updated as conditions warrant, a high-level policy should not be frequently updated.

For specific guidelines, reference the following.

- DoD Directive 5200.28, "Security Requirements for Automated Information Systems (AIS)," March 21, 1988.

- Director of Central Intelligence 1/16, "Security Policy on Intelligence Information in Automated Systems and Networks," March 14, 1988.
- <http://csrc.nist.gov/isptg/html/>, Internet Security Policy: A Technical Guide.

### 3.2.3.2 Define Information Protection System

In the Define Information Protection System activity, the user's description of information protection needs and information system environment are translated into objectives, requirements, and functions. This activity defines what the information protection system is going to do, how well the information protection system must perform its functions, and the internal and external interfaces for the information protection system.

#### 3.2.3.2.1 Information Protection Objectives

Information protection objectives have the same properties as system objectives—each will have a MoE and be unambiguous, measurable, verifiable, and traceable to an information protection need. The rationale for each objective should explain:

- The mission objectives supported by information protection objective,
- The mission-related threat driving the information protection objective,
- The consequences of not implementing the objective, and
- Information protection guidance or policy supporting the objective.

#### 3.2.3.2.2 System Context/Environment

The technical system context identifies the functions and interfaces of the system that interact with elements outside of the system boundaries. In the case of the information protection system, the mission objectives, nature of the information, mission information processing system, threats, information protection policies, and facilities strongly affect the system context. The context of the information protection system should address physical and logical boundaries between it and the mission information processing system, other systems, and the environment. Included in the context is a description of the bi-directional flow of information inputs and the outputs, signals, and energy between the system and the environment or other systems.

#### 3.2.3.2.3 Information Protection Requirements

ISSE requirements analysis activities include review and update of prior analyses (mission, threat, objectives, and system context/environment) conducted as part of the systems engineering process. As the information protection requirements evolve from the user needs to more refined system specifications, they must be sufficiently defined to permit system architecture concepts to be developed within the integrated concurrent systems engineering process. The ISSE will examine, with other information protection system stakeholders, the set of information protection requirements for correctness, completeness, coherence, interdependence, conflicts, and testability. The information protection functional, performance, interface, interoperability and derived requirements as well as design constraints will go into the RTM of the system.



### 3.2.3.2.4 Functional Analysis

The ISSE will use many of the systems engineering tools to understand the functioning and allocation of functions to various information protection CIs. The ISSE must understand how the information protection subsystem is part of and supports the overall system.

### 3.2.3.3 Design Information Protection System

In this activity, the ISSE builds the system architecture and specifies the design solution for the information protection system. As the ISSE proceeds through this activity, he or she will continue to:

- Refine, validate, and examine technical rationale for requirements and threat assessments,
- Ensure that the set of lower-level requirements satisfy system-level requirements,
- Support system-level architecture, CI, and interface definition,
- Support long lead-time and early procurement decisions,
- Define information protection verification and validation procedures and strategies,
- Consider information protection operations and life-cycle support issues,
- Continue tracking and refining information protection relevant acquisition and engineering management plans and strategies,
- Continue system-specific information protection risk reviews and assessments,
- Support the certification and accreditation processes, and
- Participate in the systems engineering process.

#### 3.2.3.3.1 Functional Allocation

As the system functions are assigned to people, hardware, software, and firmware, information protection functions are assigned to these system elements also. As functions are allocated to components, the components become responsible for satisfying the corresponding functional and performance requirements as well as a subset of the overall system constraints in the problem space. Various information protection system architectures will be examined, and the ISSE will negotiate an agreement on the information protection system architecture with system stakeholders that is both conceptually and physically feasible.

#### 3.2.3.3.2 Preliminary Information Protection Design

The entry conditions to this activity are, at a minimum, stable agreement on information protection requirements and a stable information protection system architecture under CM. Once the architecture is defined and baselined, system and ISSE engineers will generate specifications that detail what is to be built down to the CI level. Production and review of the higher level specifications occur before the PDR. ISSE activities for this activity include:

- Reviewing and refining Discover Needs and Define System activities' work products, especially definition of the CI-level and interface specifications,

- Surveying existing solutions for a match to CI-level requirements,
- Examining rationales for proposed PDR-level (of abstraction) solutions,
- Verification that CI specifications meet higher-level information protection requirements,
- Supporting the certification and accreditation processes,
- Supporting information protection operations development and life-cycle management decisions, and
- Participating in the system engineering process.

The PDR results in an Allocated System Baseline Configuration.

### 3.2.3.3.3 Detailed Information Protection Design

Detailed information protection design results in lower level product specifications that either complete the design of CIs that are under development or specify and justify the selection of CIs that are being bought. This activity will conclude with the CI-CDR — a review of each detailed CI-specification for completeness, conflicts, compatibility (with interfacing systems) verifiability, information protection risks, integration risks, and traceability to (and satisfaction of) requirements. ISSE activities for the detailed information protection system design include:

- Reviewing and refining previous Preliminary Design work products;
- Supporting system- and CI-level design by providing input on feasible information protection solutions and/or review of detailed design materials;
- Examining technical rationales for CDR-level solutions;
- Supporting, generating, and verifying information protection test and evaluation requirements and procedures;
- Tracking and applying information protection assurance mechanisms;
- Verifying CI designs meet higher level information protection requirements;
- Completing most inputs to the life-cycle security support approach, including providing information protection inputs to training and emergency training materials;
- Reviewing and updating information protection risk and threat projections as well as any changes to the requirements set;
- Supporting the certification and accreditation processes; and
- Participating in the system engineering process.

### 3.2.3.4 Implement Information Protection System

The objective of this activity is to build, buy, integrate, verify, and validate the set of CIs that will compose the information protection subsystem against the full set of information protection requirements.

The processes in this activity include those previously identified in System Engineering section. There are, however a number of additional functions that the ISSE performs in the implementation and testing of the information protection system. These include:

- Updates to the system information protection threat assessment, as projected, to the system's operational existence;
- Verification of system information protection requirements and constraints against implemented information protection solutions, and associated system verification and validation mechanisms and findings;
- Tracking of, or participation in, application of information protection assurance mechanisms related to system implementation and testing practices;
- Further inputs to and review of evolving system operational procedure and life-cycle support plans, including, for example, Communication Security (COMSEC) key distribution or releasability control issues within logistics support and information protection relevant elements within system operational and maintenance training materials;
- A formal information protection assessment in preparation for the Security Verification Review;
- Inputs to Certification and Accreditation (C&A) process activities as required; and
- Participation in the collective, multidisciplinary examination of all system issues.

These efforts and the information each produces support the Security Verification Review. Security accreditation approval would typically occur shortly after conclusion of the Security Verification Review.

### 3.2.3.4.1 Procurement

Normally, the decision to procure or produce system components is based on a hierarchy of preferred outcomes, ranging from a strong preference for Commercial-Off-The-Shelf (COTS) hardware, software, and firmware products, to a lesser preference for Government-Off-The-Shelf (GOTS) items.

A trade-off analysis is needed for a procurement/production decision. The ISSE team must ensure that the total analysis includes the relevant security factors to ensure the best overall architecture based on a balance of operation, performance, cost, schedule, and risk. In support of the decision to procure or produce system components, the ISSE team should survey the existing inventory of products to determine if there are products that satisfy the requirements for the system component. Wherever feasible, a set of potentially viable options should be identified, rather than a single source. In addition, where appropriate, the ISSE team should consider new technologies and products in ensuring the system, when implemented, will continue to be viable.

### 3.2.3.4.2 Build

In addition to the build activities detailed in the Systems Engineering section, the system design in the ISSE activity is translated into an information protection system. The purpose of this activity is to ensure that the necessary protection mechanisms have been designed and imple-

mented into the system. The information protection system, like most systems, is subjected to variables that can either enhance or degrade its effectiveness. In an information protection system, these variables can play a crucial role in determining the system's suitability for information protection. Some of these variables include:

- Physical Integrity. Have the components that are used in the production been properly safeguarded against tampering?
- Personnel Integrity. Are the people assigned to construct or assemble the system knowledgeable in proper assembly procedures, and are they cleared to the proper level necessary to ensure system trustworthiness?

As stated previously, the completion of this activity will significantly affect the remaining activities and the proper level of attention must be afforded when system assembly commences.

### 3.2.3.4.3 Test

The ISSE will have developed the information protection-related test plans and procedures. The ISSE may also have to develop test cases, tools, hardware, and software to exercise adequately the system. ISSE activities for this activity include:

- Reviewing and refining Design Information Protection System work products;
- Verifying system- and CI-level information protection requirements and constraints against implemented solutions and associated system verification and validation mechanisms and findings;
- Tracking and applying information protection assurance mechanisms related to system implementation and testing practices;
- Providing inputs to and review of the evolving life-cycle security support plans, including logistics, maintenance, and training;
- Continuing risk management activities;
- Supporting the certification and accreditation processes; and
- Participating in the systems engineering process.

### 3.2.3.5 Assess Effectiveness (Focus on Information Protection Aspects)

The ISSE focuses on the effectiveness of the information protection system. The ISSE emphasis pertains to the system's ability to provide the necessary level of confidentiality, integrity, availability, and non-repudiation to the information being processed by the system and required for mission success. If the information protection system cannot adequately meet these requirements, the success of the mission may be placed in jeopardy. This focus includes:

- Interoperability. Does the system protect information correctly across external interfaces?
- Availability. Is the system available to users to protect information and information assets?

- Training. What degree of instruction is required for users to be qualified to operate and maintain the information protection system?
- Human/Machine Interface. Does the human/machine interface contribute to users making mistakes or compromising information protection mechanisms?
- Cost. Is it financially feasible to construct and/or maintain the information protection system?

## 3.3 Relation of ISSE Process to Other Processes

### 3.3.1 Purpose and Outline of 3.3

In the following sections several commonly used system processes are presented and compared to the ISSE process explained in Section 3.2 (Information Systems Security Engineering). The system processes include systems acquisition, risk management, life-cycle support, certification and accreditation (specifically the DoD Information Technology Security Certification and Accreditation Process (DITSCAP)), and the specification process in the Common Criteria (CC). The relationship of each system process to the individual activities of the ISSE process is explained in this section. The basic activities of the ISSE process include discovering the mission needs, defining the system functional characteristics, designing the system to provide the identified functionality, and implementing the system in accordance with the design. All of the processes contain similar activities that can be tailored to meet a specific developmental need or requirement. The commonality among the system processes and ISSE enables developers to simultaneously complete developmental activities that are required for both an information protection process and a life-cycle process. This reduces the need to duplicate system development efforts and decreases the time in which a system or product can be securely developed and placed into production. The first ISSE process comparison is performed against a systems acquisition process, followed by a comparison to a risk management process, a life-cycle support process, a certification and accreditation process (specifically the DITSCAP), and finally the Common Criteria.

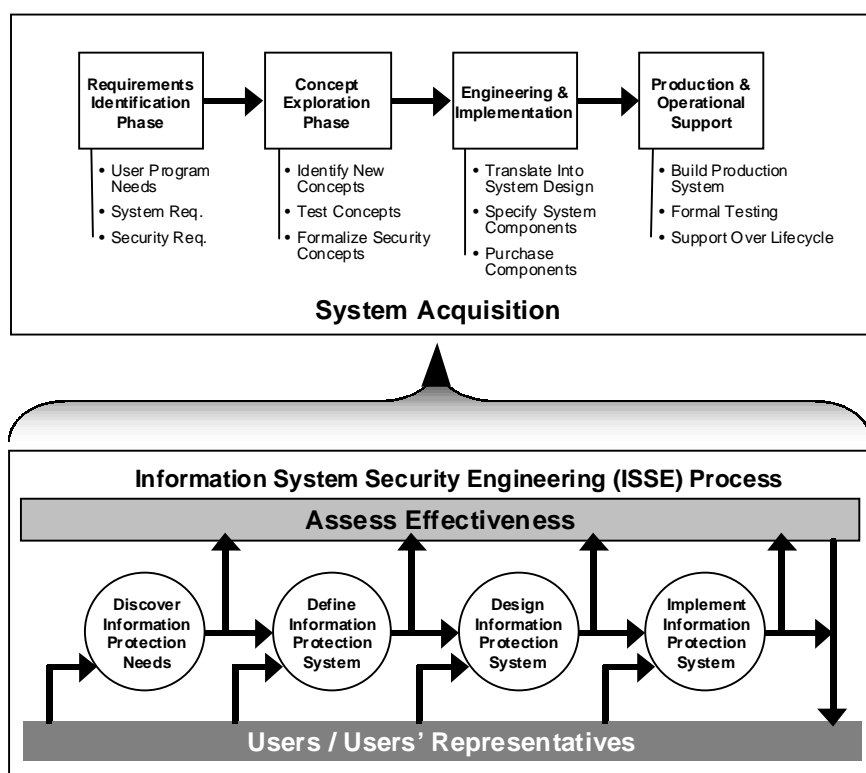
### 3.3.2 Systems Acquisition Process

An underlying process within any system engineering approach is the system acquisition process. The system acquisition process ideally occurs in parallel with the systems engineering and ISSE processes. System acquisition phases include a requirements identification phase, a concept exploration phase, an engineering and implementation phase, and a production and operational support phase. Figure 3-4 and the following paragraphs identify acquisition phases and the corresponding ISSE activities.

At this point, the ISSE understands the mission needs, the mission environment, the policy constraints, and the system information protection needs. Now a judgement must be made concerning whether, given this information and certain time, cost, technological, and risk estimates

for an information protection subsystem, an information protection subsystem satisfying these needs is feasible.

Before proceeding, someone in authority, such as the Designated Approval Authority (DAA), must make a risk-based decision that the information protection system concept is feasible given the system acquisition and engineering management situation. The decision to proceed or not should be documented so that if circumstances change, a reasoned decision reconsideration can be performed without redoing all of the work that went into the initial decision.



**Figure 3-4 ISSE and System Acquisition Process Flows**

During the acquisition requirements identification phase (which parallels the systems engineering (and ISSE) Discover (Information Protection) System activities), acquisition activities will include reviewing a program's needs and defining system information protection requirements from the user's point of view. Exploitation of new technologies will be investigated, potential threats to system operation will be identified, and alternative solutions will be investigated. Cost studies will be performed, risk assessments will be developed, technical feasibility will be determined, and life-cycle cost estimates will be developed. Finally, a decision to proceed to the next phase will be made and funding to continue the development process will be requested.

During the acquisition concept exploration phase (which parallels the systems engineering (and ISSE) Define (Information Protection) System activities), acquisition activities that explore new concepts and solidify system definition will occur. One or more promising alternatives will be selected for proof of concept testing. Functional, performance, and information protection characteristics will be identified. Systems engineering plans and life-cycle support plans will

be developed to support system acquisition. Simulation, prototyping, and testing will be used to prove design concepts in an effort to formalize information protection concepts into specific design requirements and system-level specifications that meet the stakeholders needs. The system-level specifications and formalized information protection concepts defined in this phase will be used to design the system.

During the acquisition engineering and implementation phase (which parallels the systems engineering (and ISSE)) Design (Information Protection) System activities), acquisition activities will focus on ensuring that previous developments and prototypes are translated into a system design that is stable, producible, and cost-effective. Design work will be completed, formal specifications will be released, and specified components will be either built or made. To support acquisition efforts, system components will be identified and functionality characteristics will be documented to aid with component acquisition.

During the acquisition production and operational support phase (which parallels the systems engineering (and ISSE) Implement (Information Protection) System activities), acquisition activities include integrating system components and delivering the system to the stakeholders. Formal testing to demonstrate that system requirements have been met will be conducted. Continued support of the system will be required over the life-cycle of the system. Continuing customer training and support will be required. Spare parts inventories will need to be produced and maintained along with patches or Engineering Change Proposals (ECP) to fix system design flaws.

There are two special cases that Figure 3-4 does not show. Often the system acquisition process gets ahead of the ISSE process. When the system architecture is defined and components have been selected before the mission information protection needs have been defined, then, if the information protection requirements are to be satisfied, the systems acquisition process will have to be reversed and returned to the requirements identification phase. At this point, to satisfy the “new” information protection requirements, the ISSE and systems engineer will have to satisfy many more constraints. Namely, the design constraints imposed by the parts of the system already defined that will not change, for whatever reason, must be satisfied. This is a very expensive way to build a sub-optimal system.

The other case not explicitly shown in Figure 3-4 is the case of the system upgrade. In this case, the system is already operational. The new requirements for the system will come from a discovered capability shortfall or new mission statement. There will be interface requirements because the new system components must interact with and not degrade the overall system health. System upgrades should go through the ISSE and system acquisition processes starting with the performance and functional requirements while taking note of the increased interface requirements from the existing system. Depending on the extensiveness of the upgrade, it may be treated almost as a new development with some pre-selected pieces from the old system.

Again, it is possible to recognize that common process elements do exist between the system acquisition process and the ISSE process. The process similarities show that it is possible to incorporate information protection design concepts into the system design processes with very little process modification required.

### 3.3.3 Risk Management

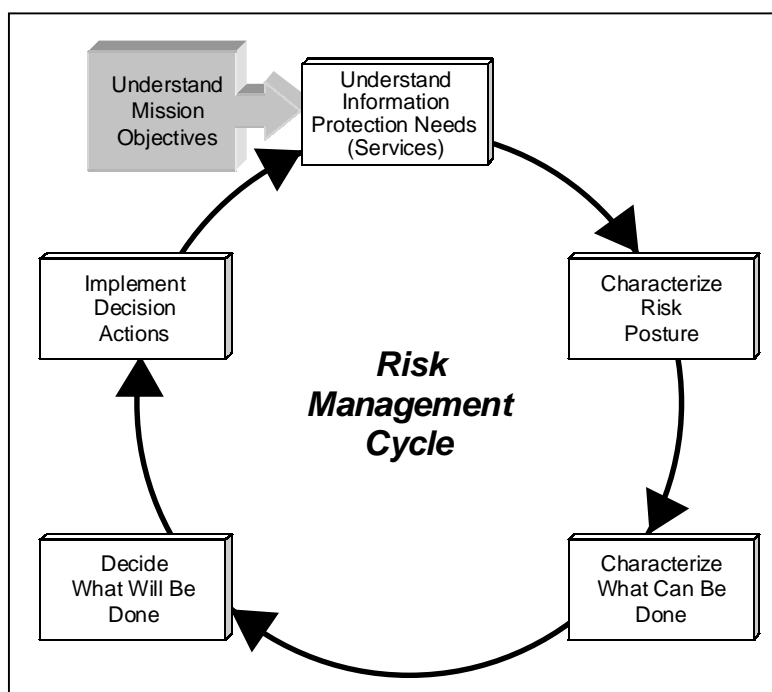
Risk management is applied during the initial system development, throughout the development and acquisition process, and on existing systems. During the initial system acquisition process, risk management activities are often the same as activities performed during the ISSE process. Once the system is fielded, the risk management process needs to be able to adjust and respond to changes in the system design and configuration, changes in the operating environment, and changes to the supported mission—those elements that cause a resulting change in risk. As these parameters change over time throughout development, acquisition, and fielding, it is necessary to periodically revisit the understanding of the risks incurred operating within the current and projected environment and determine if a change in protection approach (technology, procedures, personnel) is warranted. These decisions concerning changes and their ultimate implementation are the focus of the process.

Risk management is cyclical in nature and starts with understanding the mission and information protection objectives and needs, characterizing the current risk posture of the mission, and characterizing the possible impediments to successful mission completion. Successful mission completion may require developing, designing, integrating, and deploying an information protection system that has certain technical (operational and information protection), cost, and schedule (time to develop) features as well as an expected useful lifetime. After understanding the mission (and information protection needs), the next step is to identify, characterize, and prioritize the risks. Now, the ISSE is in a position to recommend to management countermeasure options to mitigate the risks, which will aid the ultimate decision maker in deciding which risks can and should be mitigated. For a given user's system, usually it is not possible to implement a set of technical and non-technical countermeasures that completely satisfy the user's organizational information protection policy and mitigate all risks. The nature of the mission may require operational capabilities and functionality that is incompatible with a risk minimization posture. Therefore, a risk management methodology should be applied in deciding whether to field a specific information protection system. The first step in such a strategy is to assess thoroughly the information protection risks associated with a given system. The residual risks to a system are then based on the level of probability or the likelihood that an adversary will successfully perform a particular attack that will affect the operational mission.

The ISSE will go through the following cyclical process (Figure 3-5) to build the framework for justifying subsequent decisions:

- Understand Mission Objectives;
- Understand Information Protection Needs;
- Characterize the Risk Posture;
- Characterize What Can Be Done ;
- Decide What Will Be Done;
- Implement Decision Actions.





**Figure 3-5 Risk Management Process**

Figure 3-5 and the following subsections address the steps required to identify and assess information systems information protection engineering risks.

### 3.3.3.1 Understand Mission and Information Protection Objectives

The risk management process starts with the determination of the user's mission objectives and the mission's information protection needs. This is done in parallel with the systems engineer by identifying the mission of the organization, its mission criticality, the number of users, the operating environment, and the user's current information protection posture. It requires a review of the relevant policies, regulations, and applicable operating guidelines and standards.

A review of the user's operating procedures and current system architecture is part of the collective effort to correctly value the information and the information protection needs. The information that is garnered from this step in the process will enable the ISSE to determine the potential risks that will be introduced with the implementation of the system and the level of protection needed.

### 3.3.3.2 Characterize Risk Posture

In this activity of risk management the ISSE is focused on gaining insight into the risks that are being incurred or will be incurred based on system operation and design and the hostile operating environment. This activity includes the supporting risk, vulnerability/attack, threat, and

mission impact analyses. These are then synthesized to provide analytical insights to the pressing operational and budgetary questions pertinent to the decision maker.

### 3.3.3.2.1 Risk Analysis

Risk analysis is that portion of risk management that collects and analyzes risk-related data to provide decision makers information on the benefits and costs of alternative courses of action for executing the assigned mission in a hostile environment. These are difficult decisions that require the ISSE to conduct effective analysis. Figure 3-6 diagrams a basic synthesis of information that helps determine the costs and benefits of various alternative courses of actions.

When performing a risk analysis the following factors should be considered:

- Threats. Threat types can be broken down into three distinct groups, adversarial, non-adversarial, and natural.
  - Adversarial. Each adversarial relationship is unique and must be defined and analyzed. Adversarial relationships depend on the situation. In the case of foreign states, an adversary today may be an ally tomorrow or vice versa. Examples of potential adversaries are listed below:
    - Terrorists
    - Unbalanced persons
    - Criminals - alone
    - Criminals - organized
    - Insider: in collusion with outsider
    - Insider: in collusion with outsider
    - Insider: disgruntled, unbalanced, spy, criminal

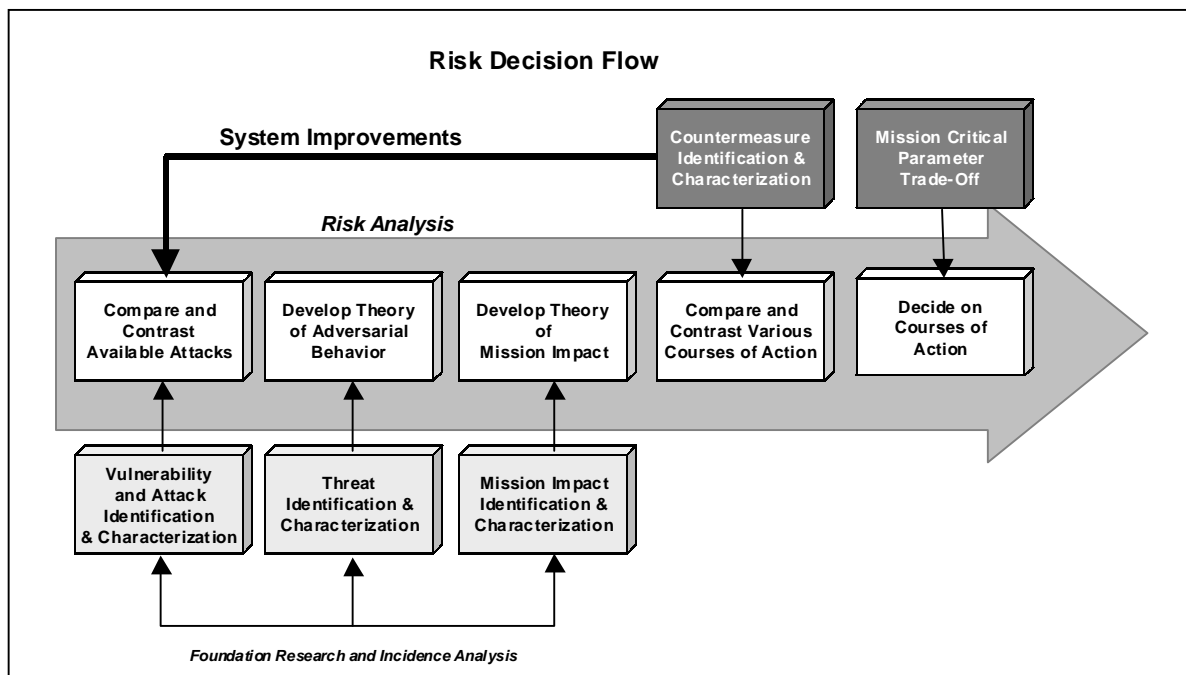


Figure 3-6 Risk Decision Flow

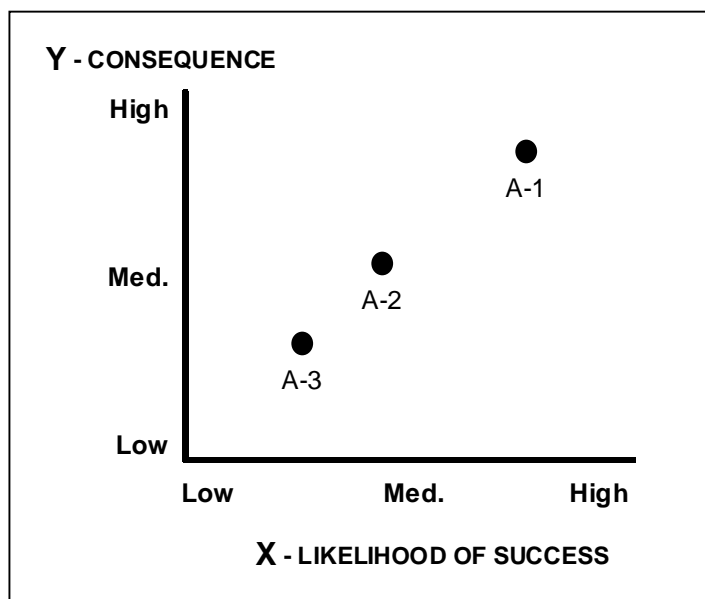
- Non-adversarial. These types of threats have no malicious objectives, motivations, or intentions. However, they do have the capability to create a degree of harm that sometimes surpasses that of an adversary with intent to harm. Having documented information protection policies that are enforced will greatly reduce the risk posed by this type of threat element. Examples of potential non-adversarial threats are listed below:
  - System users, and
  - Maintenance personnel.
- Natural. These threats may be predicted for any geolocation based on historical data. We have a good idea of how often they might happen and in what intensity. As with non-adversarial threats, we usually will not know exactly when this type of threat may occur. Examples of natural threats:
  - Earthquakes,
  - Volcanoes,
  - Hurricanes,
  - Tornadoes,
  - Floods,
  - Lightning,
  - Hail.
- Cost. For every attack that is launched by an adversary against a system there are associated costs that the attacker must incur. These costs can vary in size based on the level of protection of the system being attacked, and the level of or degree of sophistication of the attack itself. From a costing aspect, attacks can be divided into the following types:
  - Hardware. The cost can be expressed in terms of equipment cost to perform the analysis.
  - Software/Attack Tools. The costs associated with assembling a suite of applications or utilities that enable the attacker to launch the attack. The tools can range in complexity and cost and can be, in some cases, available on the Internet.
  - Signal Analysis. The cost can be expressed in terms of equipment used to circumvent security rules by using password discovery schemes and other methods that can be used to bypass conventional information protection policy enforcement routines.
  - Expertise. Any attack will require people at some level of expertise to develop the scheme and tools required. The cost can be expressed in terms of level of sophistication and organization required.
  - Access. Every attack requires some level of access. The cost can be expressed in terms of an attack being initiated by an insider or outsider.
- Countermeasures. The ISSE must identify what possible countermeasures can be implemented to eliminate or reduce the risk of a successful attack being launched against the system or systems under review. Unfortunately, there are usually insufficient resources to pursue this approach. The limited, available resources need to be more efficiently expended against the vulnerabilities and attacks that will be most harmful to the mission. The ISSE then must provide as much detailed information as possible to the decision makers who ultimately must decide what level of risk can be tolerated.

- **Determining the Likelihood of an Attack Occurring.** The ISSE must determine the likelihood of an attack occurring—the probability that the adversary has the motivation and capability to successfully exploit a system vulnerability.
- **Recovery from Attack.** For each attack, the ISSE must determine whether there is an ability to recover from the attack and the time (from the point of attack detection) required to correct the effects of the attack.
- **Determine Consequences to Operational Mission** – Initially, the user organization works with the “search for vulnerabilities” team to translate the technical consequences of the attacks to the organization’s operational mission; the recovery from attack aspect should also be considered in determining the consequences of an attack. A relative (high-to-low) level of consequence is identified for the individual attacks

For each attack, the ISSE will generate a risk plane or other appropriate analytical tool to show the likelihood of occurrence and its consequences to the operational mission. The risk plane is input into the certification process described in the next section. Figure 3-7 shows a typical risk plane.

### 3.3.3.3 Characterize What Can Be Done

The ISSE, after completing the risk analysis, will use the synthesized information and provide the decision makers with courses of action. The ISSE will need to compare and contrast each of the available courses of action. This analysis will provide an insight into how various countermeasure response actions improve or degrade the current risk posture and the mission capability of the system. To characterize what can be done, the ISSE will need to focus on developing sets of possible courses of action along with their relative costs and benefits. These options can include several different possible solutions to include the “no change” option, the “shut down” option, and various combinations of technical, procedural, and personnel changes to the system to mitigate potential attacks, reduce the operational impact of successful attacks, or change the operational capability/information protection risk balance. The ISSE will also conduct a cost-benefit analysis to outline clearly the degree of protection that can be provided, given the cost of countermeasures.



**Figure 3-7 Risk Plane**

### 3.3.3.4 Decide What Will Be Done

Based on the risk analysis, the decision maker will have to decide which threats and vulnerabilities will be countered. The decision will be based on the threat/vulnerability risk, the cost

of countermeasures in dollars and mission functionality, and the effectiveness of the countermeasures in mitigating the overall risk. The ISSE will have to document what the residual risks are from threats/vulnerabilities that are not countered at all or that are incompletely countered.

### 3.3.3.5 Implement Decision Actions

Implementation of an information protection system is complete when the system is operating at a level of risk acceptable to the local DAA. This is accomplished in the two activities of post-development:

- Verification and
- Validation.

Verification is the process of ensuring that implementation of the system satisfies the requirements as stated and documented. In short, verification answers the question, “Did you build the system right?” Validation re-verifies that stated objectives are met and that the system is ready for full implementation. In short, validation answers the question “Did you build the right system?”

A certification analysis is performed to ascertain the ability of the system to comply with the information protection requirements and minimize the residual risk. Information protection verification testing includes cryptographic verification testing and functional information protection testing. It ensures that all the information protection functions are implemented exactly as described in the specification. Information protection testing searches for unspecified, but implemented, functions that may subvert information protection critical operations and any vulnerabilities resulting from improper implementations of information protection features.

#### 3.3.3.5.1 Assess Outcome of Decision Actions

To complete the cycle of risk management, the ISSE must assess the outcomes of the decisions made. Not only does the ISSE have to re-evaluate his or her own decisions, but also the context in which the decisions were made. If new threats, vulnerabilities, or technologies are discovered that change the attack/defend balance, then the ISSE must re-evaluate the risks. If a political change affects the likelihood of attack or the motivations of the attackers, this is also reason for re-examination of the risks. A change in mission needs may force a change in the mission capability/information protection risk balance. Risk Management is a continuous process that provides input into every systems engineering and ISSE activity and affects every stage of the system life-cycle.

### 3.3.4 Life-Cycle Support

In general, a system life-cycle consists of the following stages: concept and requirements definition; system functional design; system development and acquisition; system implementation and test; system sustained operational support; and finally system disposal. The approach to system life-cycle support has transformed over the past several years in response to the need to incorporate security components and processes into systems engineering processes. As a result,

the ISSE methodology described in Section 3.2 (Information Systems Security Engineering ) was developed as an example of an information security life-cycle process.

As with any functional or system-level requirement, security should be addressed early in the life-cycle process in terms of understanding the security requirements, participating in the evaluation of security products, and finally in the engineering, design, and implementation of the system. Over the years, lessons learned have shown that it is very difficult to retrofit security solutions after the system has been developed. Therefore, security requirements must be considered while discovering the needs and defining the system. To effectively integrate security measures and controls into the systems engineering processes, designers and developers had to modify existing process models to incorporate an iterative system development life-cycle that focuses on security controls and protection mechanisms needed to secure a system. A successful example of this philosophy is the DITSCAP.

Although the DITSCAP is called a certification and accreditation process, it is based heavily on ISSE concepts and life-cycle management principles. The phases of the DITSCAP can be correlated to the basic ISSE activities identified above. In the DITSCAP, the System Security Authorization Agreement (SSAA) is used to document all security criteria used throughout the Information Technology (IT) system life-cycle. The majority of the system life-cycle analysis takes place in Phase 2 of the DITSCAP. During Phase 2, life-cycle plans are developed and evaluated. Life-cycle support activities include:

- Creation of a life-cycle management plan,
- Creation of a system engineering management plan,
- Creation of configuration control procedures,
- Creation of security plans,
- Creation of maintenance plans,
- Creation of contingency and continuity of operation plans, and
- Creation of system disposal plans.

Whether using the DITSCAP, a systems engineering process, or a system life-cycle process, it is possible to recognize that common process elements do exist. Because of the cross process commonality, it is possible to ensure that security components are incorporated into all system development activities with minimal cost, schedule, or functional impact.

The compliance validation process is focused on management of the periodic re-accreditation of systems and the incorporation of a variety of methods that are responsive to managed levels of risk within the deployed systems. Traditional inspection methods are used, but a managed approach based on the availability of “living documentation” and the inspection of previous systems weigh in the inspection method selected. Additionally, if a system is found to have a more limited risk posture based on less significant changes, then the attendant inspection process will be appropriately simplified.

The final step in the compliance revalidation is an inspection. The purpose of inspection tailoring is to develop inspection procedures that focus on the areas that are most vulnerable. The method and level of detail of an inspection are tailored according to risk classification and particular areas of risk as determined in a formal risk analysis.

### 3.3.5 Certification and Accreditation

Certification and accreditation (C&A) is defined as a comprehensive evaluation of the technical and non-technical security features of an AIS and other safeguards to establish the extent to which a particular design and implementation meets a set of specified security requirements and the formal declaration by a DAA that the information system is approved to operate at an acceptable level of risk. The Certification Authority (CA) is the official responsible for performing the evaluation of these security features and other safeguards, made in support of the accreditation process. This person establishes the extent that a particular design and implementation meet a set of specified security requirements. The initial certification tasks include:

- System architecture analysis,
- Software design analysis,
- Network connection rule compliance analysis,
- Integrity analysis of integrated products,
- Life-cycle management analysis, and
- Vulnerability assessment.

A system undergoing C&A should have formal documentation supporting each task, a documented security specification, a comprehensive test plan, and written assurance that all network and other interconnection requirements have been implemented. The specific certification tasks may be tailored to the system's program strategy, its life-cycle management process, and the position of the information system in its life-cycle. Certification tasks are tailored to the system development activities to ensure that the former are relevant to the process and provide the required degree of analysis to ensure conformance with written documentation, usually in the form of an SSAA.

The vulnerability assessment evaluates security vulnerabilities concerning confidentiality, integrity, availability, and accountability and recommends applicable countermeasures. The DAA should determine the acceptable level of risk to protect the system commensurate with its value to the DoD. The vulnerability assessment concentrates on the progress in implementing the security requirements of the SSAA. The assessment conclusions are used to determine that the AIS is ready for formal C&A evaluation and testing. Final certification tasks include:

- Security test and evaluation,
- Penetration testing,
- TEMPEST and Red-Black verification,
- Validation of COMSEC compliance,
- System management analysis,
- Site accreditation survey,
- Contingency plan evaluation, and
- Risk-based management review.

If the CA concludes that the information system satisfies the SSAA technical requirements, the CA issues a system certification. That is a certification that the information system has com-

plied with the agreed-on security requirements. Supplemental recommendations also might be made to improve the system's security posture. Such recommendations should also provide input to future system enhancements and change management decisions.

The CA's recommendation, the DAA authorization to operate, the supporting documentation, and the SSAA form the accreditation package. The supporting documentation may vary between system classes. That documentation, at minimum, should include security findings, deficiencies, and risks of operation. The accreditation package must contain all information necessary to support the recommended decision. If the decision is to accredit, the decision should include the security parameters under which the information system in its computing environment is authorized to operate. If the system does not meet the requirements stated in the SSAA, but mission criticality mandates that the system become operational, a temporary approval may be issued.

Final accreditations are good for three years or until a significant change warrants a new accreditation. Systems requiring revalidation are identified based on a request notification, systems that require a periodic re-evaluation, or systems that have exhibited characteristics that indicate a security problem exists. Periodically scheduled systems are default candidates, while behavior related identification is based on a triggering event

### 3.3.5.1 DITSCAP

The DITSCAP drives the DoD C&A process. It establishes the standard process for C&A for IT systems in DoD. For intelligence systems, Director of Central Intelligence Directive (DCID) 6/3 drives the C&A process. The objective of the DITSCAP is to establish a DoD standard infrastructure-centric approach that protects and secures the entities composing the Defense Information Infrastructure (DII). The set of activities presented in the DITSCAP standardizes the C&A process for single IT entities and leads to both a more secure system operation and a more secure DII. The process considers the system mission, environment, and architecture while assessing the impact of operation of that system on the DII. The process:

- Implements policy, assigns responsibilities, and prescribes procedures for C&A of IT systems, including AIS, networks, and sites in the DoD.
- Creates the DITSCAP for information protection C&A of unclassified and classified IT systems to implement reference 1.
- Stresses the importance of a life-cycle management approach to the C&A and re-accreditation of DoD IT.

The DITSCAP applies to all systems requiring C&A throughout their life-cycle. It is designed to be adaptable to any type of IT system and any computing environment and mission. It may be adapted to include existing system certifications, evaluated products, use new security technology or programs, and adjust to the applicable standards. The DITSCAP may be mapped to any system life-cycle process but is independent of the life-cycle strategy. The DITSCAP is designed to adjust to the development, modification, and operational life-cycle stages. Each new C&A effort begins with Phase 1, Definition, and ends with Phase 4, Post Accreditation, in which follow-up actions ensure that the approved information system or system component continues to operate in its computing environment in accordance with its accreditation.



DITSCAP is separated into a logical sequence of phases that will ultimately lead the system to its final accreditation. Figure 3-8 presents the phases:

- Phase 1 - Definition,
- Phase 2 - Verification,
- Phase 3 - Validation, and
- Phase 4 - Post Accreditation.

Phase 1 of the DITSCAP, Definition, concludes with the written SSAA—an agreement between the Program Manager, the DAA, the CA, and the Users Representative on the description of system mission, target environment, target architecture, and security policies. The SSAA describes the planning and certification actions, resources, and documentation required to support certification and accreditation. The specific roles of the signature authorities are explained in reference 1.

In the ISSE process, the flow from the system security requirements goes to a defined system architecture. Both the ISSE Define Information Protection System and DITSCAP Phase 2 activities, Verification, verify that the system security functions and architecture meet the system functional requirements for information processing and security. The Life-cycle Management activity requires documentation of many of the systems engineering documents used to direct, control, and manage the development, design, operation, and disposal of the system. The results of these activities may or may not result in a change to the initial SSAA, system functions, and system architecture. At this point, the system security engineer will have developed test plans and procedures that cover system functions, system integration, and security compliance. The system security engineer should evaluate and describe the level of residual risk remaining as a result of the proposed design.

The activities in Phase 3 of the DITSCAP, Validation, bring the system into existence, verify that the design of the system is correct, and validate that the system meets the requirements. This phase of activity culminates in a system that is verified, integrated, validated, certified, and accredited. Descriptions of the various DITSCAP Phase 3 activities are provided in reference 1.

Post Accreditation, Phase 4 of the DITSCAP, deals with the Compliance Validation (CV) process. The Compliance Validation Inspection (CVI) process is focused on management of the periodic re-accreditation of systems and the incorporation of a variety of methods that are responsive to managed levels of risk within the deployed systems. Traditional inspection methods are used, but a managed approach based on the availability of “living documentation” and the previous inspection of systems according to Phase 3 of the DITSCAP, weigh in the inspection method selected. Additionally, if a system is found to have a more limited risk posture based on less significant changes, the attendant inspection process will be appropriately simplified.

The supplemental DITSCAP document describes compliance validation in terms of the following processes: Process Management, Risk Management, CM Review, CVI Prioritization, and Inspection Tailoring. The intent of these processes is to apply resources effectively toward the CVI process, particularly when multiple systems are presented for revalidation or when behavior dictates the need for CV. The processes can be used for stand-alone or multiple systems; however, using processes for multiple systems is more effective.

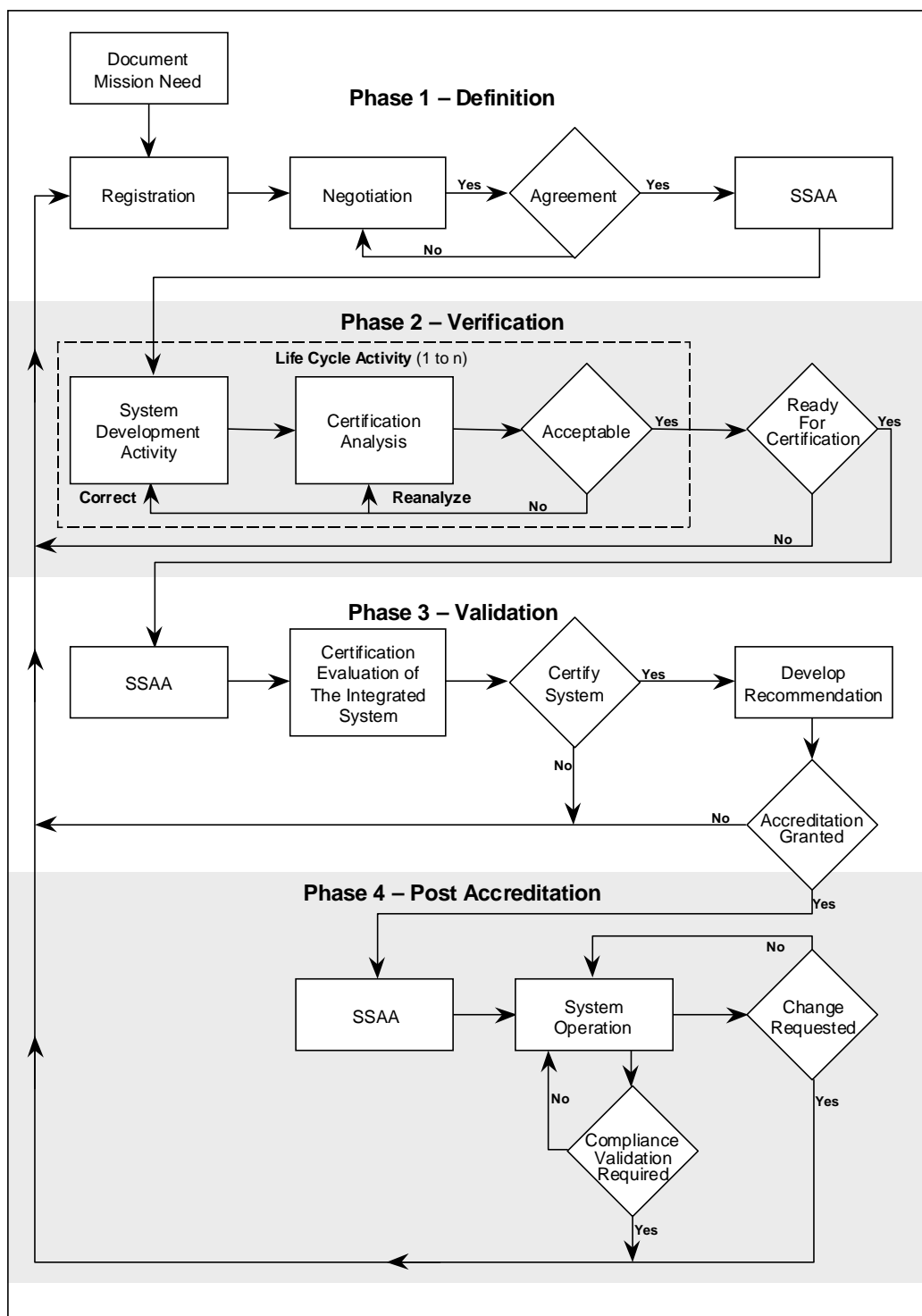
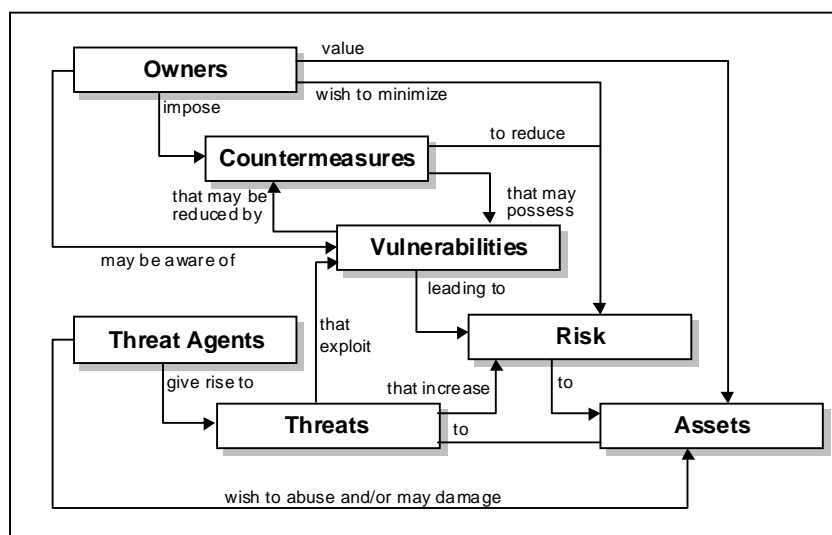


Figure 3-8 DITSCAP Flow

### 3.3.6 Common Criteria and Possible Uses

The Common Criteria (CC) is an international standard that defines criteria “... to be used as the basis for evaluation of security properties of IT products and systems ... . This standard will permit comparability between the results of independent security evaluations ... by providing a common set of requirements for the security functions of IT products and systems and for assurance measures applied to them during a security evaluation. The evaluation process establishes a level of confidence that the security functions of such products and systems and the assurance measures applied to them meet these requirements.” The CC comprises three main parts—the Security Functional Components, the Security Assurance Components, and the Evaluation Method. The ISSE may use the CC as a tool to support ISSE activities up to and including developing the system-level specifications for the information protection system and to support accreditation.

Figure 3-9 shows how the CC is meant to be used. The process for building information security in the grammar of the Common Criteria follows the ISSE process. The Discover Information Protection Needs activity will provide the information required to describe how owners value assets, what are the threat agents, what are the threats, what are the countermeasures (requirements and functions), and what are the risks (partially). The Define Information Protection System activity provides information to describe what are the countermeasures (named components), what are the vulnerabilities (based on the architecture), what are the risks (more fully). The Design Information Protection System activity provides information to describe what are the countermeasures (verified information protection product functions), what are the vulnerabilities (based on the design, and unit and verification test results) and what are the risks (more fully). The Implement Information Protection System activity finally provides information to describe what are the countermeasures (installed, validated information protection system functions), what are the vulnerabilities (based on the validation and penetration test results), and what are the risks (more fully). The CC does not address personnel or operational security, but these security measures must be discussed as part of the secure environment. The CC also does not address Assess Effectiveness or other management practices required to bring a system into being effectively. The CC provides a standard language and grammar that customers and developers can use to make claims of general performance (Protection Profile or PP) or claims of specific performance (Security Target or ST)



**Figure 3-9 Security Concepts and Relationships in the Common Criteria**

The CC does not address personnel or operational security, but these security measures must be discussed as part of the secure environment. The CC also does not address Assess Effectiveness or other management practices required to bring a system into being effectively. The CC provides a standard language and grammar that customers and developers can use to make claims of general performance (Protection Profile or PP) or claims of specific performance (Security Target or ST)

that may be evaluated. Note that what the Common Criteria calls a threat that exploits a vulnerability is called an attack in subsequent chapters of the IATF.

A PP defines, in a standardized form, a set of functional and assurance security requirements either from the CC or user-defined that are needed to address a known or assumed security problem (which may be defined in terms of identified threats to assets requiring protection). The PP permits the implementation independent expression of security requirements for a set of Targets of Evaluation (TOEs) that will comply fully with a set of security objectives. A PP is intended to be reusable and to define TOE requirements that are known to be useful and effective in meeting the identified objectives, both for functions and assurance. A PP also contains the rationale for security objectives and security. When the (TOE) is a particular type or class of IT product or system (e.g., operating system, database management system, smartcard, firewall, etc.), the definition of security requirements will not be specific to any particular system or product.

A PP could be developed by user communities, IT product developers, or other parties interested in defining such a common set of requirements. A PP gives consumers a means of referring to a specific set of security needs and facilitates future evaluation against those needs. As such, a PP is an appropriate CC document for the ISSE to develop to describe an architecture, possibly as a basis for acquisition and technical evaluation.

An ST contains a set of security requirements that may be made by reference to a PP, directly by reference to CC functional or assurance components, or stated explicitly. An ST permits the expression of security requirements for a specific TOE that are shown, by evaluation, to be useful and effective in meeting the identified objectives. An ST contains the TOE summary specification, together with the security requirements and objectives, and the rationale for each. An ST is the basis for agreement between all parties as to what security the TOE offers. As such, the ST is an appropriate CC document to develop to describe a system-level specification, possibly as a basis for acquisition and support for accreditation.

The PP and ST can also be used as a means of communication among the parties responsible for managing the development of a system, the stakeholders in that system, and the organization responsible for producing the system. In this environment, the ST is proposed in response to the PP. The content of the PP and ST may be negotiated among the players. Evaluation of the actual system against the PP and ST may be part of the acceptance process. In general, non-IT security requirements will also be negotiated and evaluated. Often the solution of the security problem will not be independent of other system requirements. The relationship between STs and PPs is shown Figure 3-10.

The CC philosophy is to provide assurance based on an evaluation (active investigation) of the IT product or system that is to be trusted. Evaluation has been the traditional means of providing assurance and is the basis for prior evaluation criteria documents. In aligning the existing approaches, the CC adopts the same philosophy. The CC proposes measuring the validity of the documentation and of the resulting IT product or system by expert evaluators with increasing emphasis on scope, depth, and rigor.

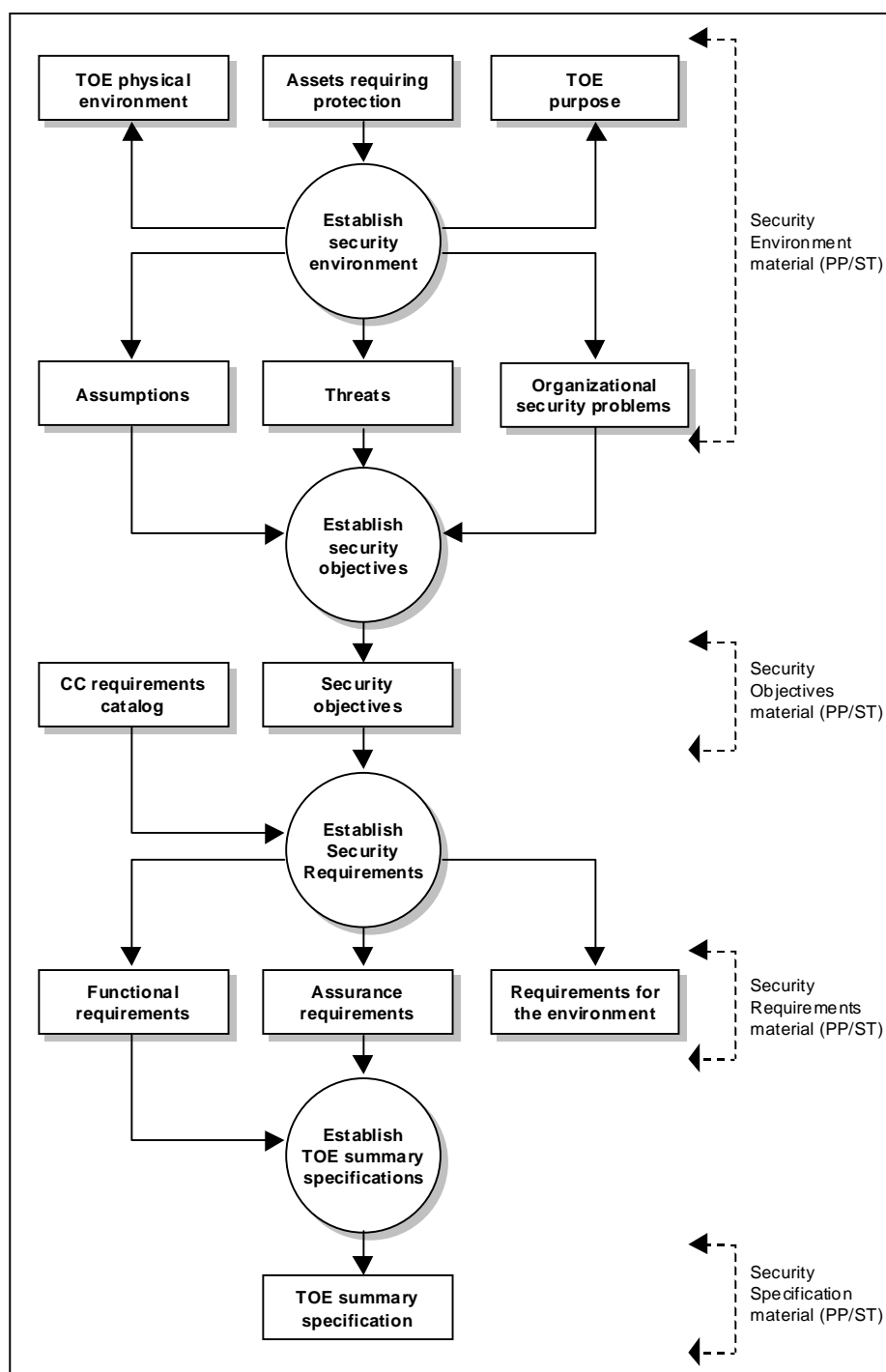
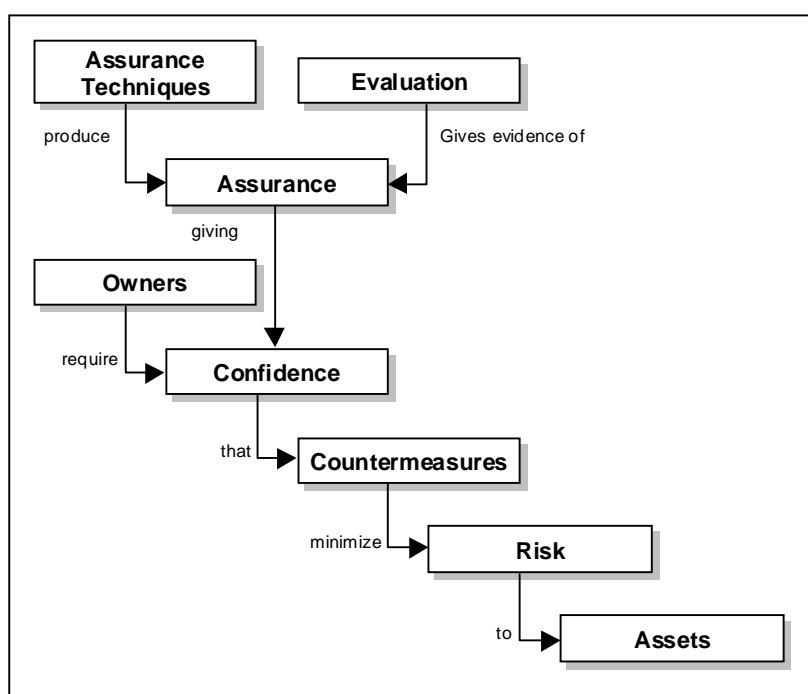


Figure 3-10 Protection Profile — Security Target Relationship

The CC does not exclude, nor does it comment on, the relative merits of other means of gaining assurance. Research continues with respect to alternative ways of gaining assurance. As mature alternative approaches emerge from these research activities, they will be considered for inclusion in the Common Criteria, which is so structured as to allow their future introduction.

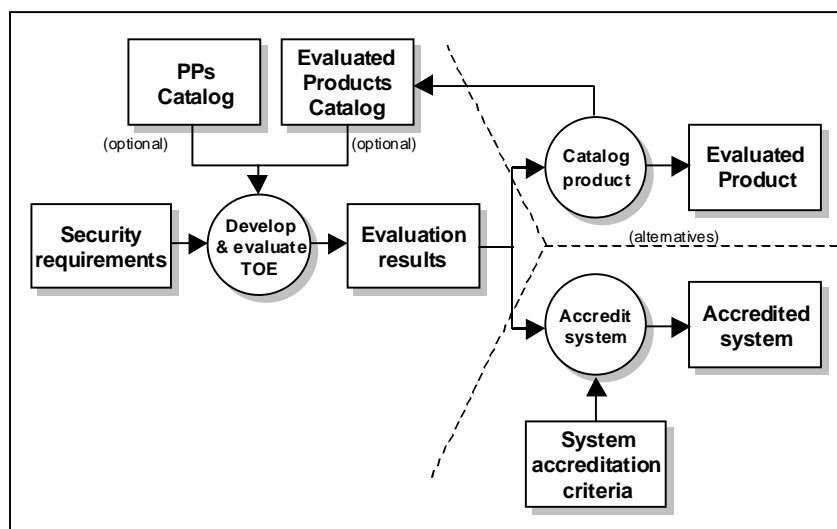
The CC philosophy asserts that greater assurance results from the application of greater evaluation effort, and the goal is to apply the minimum effort required to provide the necessary level of assurance. The increasing level of effort is based on:

- Scope. The effort is greater because a larger portion of the IT product or system is included.
- Depth. The effort is greater because evaluation evidence gathering is engaged at a finer level of design and implementation detail.
- Rigor. The effort is greater because evaluation evidence gathering is applied in a more structured, formal manner.
- The evaluation process provides evidence of the assurances required in the PP or ST as shown in Figure 3-11. The result of evaluation is some level of confidence in the information protection system. Other ISSE processes, such as risk management or DITSCAP, provide the means for translating the confidence in countermeasures (architecture or design, as appropriate to the development cycle) into management decision criteria.



**Figure 3-11 Evaluation Concepts and Relationships**

Figure 3-12 shows that systems (or subsystems) may be evaluated against PPs or STs to create a catalog of evaluated products or, with outside system accreditation criteria (from DITSCAP, for example), to support accrediting a system.



**Figure 3-12 Uses of Evaluation Results**

## References

1. Merriam-Webster's Collegiate Dictionary, <http://www.m-w.com/>.
2. DoD Directive 5200.40, *DoD Information Technology Security Certification and Accreditation Process (DITSCAP)*, November 30, 1997.
3. SABI Risk Team. *Basic Risk Management for DoD Information Systems*, Edition 1.1, 21 January 2000.
4. ND186. *Introduction to Information Systems Security Engineering (ISSE)*, Session 02-99, May 1999.

## Additional References

- a. ND208. Fundamentals of the Risk Management Process, 1999.
- b. Information Systems Security Engineering Handbook, Release 1.0, 28 February 1994.
- c. Information System Security Policy Guideline, I942-TR-003, 1994.
- d. Common Criteria for Information Technology Security Evaluation, Part 1: Introduction and general model, Version 2.1, CCIMB-99-031, August 1999.
- e. Common Criteria for Information Technology Security Evaluation, Part 2: Security Functional Requirements, Version 2.1, CCIMB-99-032, August 1999.
- f. Common Criteria for Information Technology Security Evaluation, Part 3: Security Assurance Requirements, Version 2.1, CCIMB-99-033, August 1999.



## Chapter 4

# Technical Security Countermeasures

The authors of the Information Assurance Technical Framework IATF recognize the importance of using both technical and non-technical countermeasures in formulating an effective overall security solution to address threats at all layers of the information infrastructure. This chapter of the IATF presents a general discussion of the principles of determining appropriate technical security countermeasures. It includes a detailed assessment on threats of, important security services, robustness strategy, interoperability framework, and the Key Management Infrastructure (KMI)/Public Key Infrastructure (PKI). It also provides background for the detailed technical discussions contained in later sections of the IATF.

## 4.1 Introduction

Adversaries' primary goals can be grouped into three general categories: unauthorized access, unauthorized modification, and denial of authorized access. The intent of a security solution is to prevent an adversary from successfully achieving these goals. This chapter discusses threats, security services and appropriate security technologies. Using the methodology described in Chapter 3 (Information Systems Security Methodology), with the consideration of applicable threats, security solutions may be proposed which support appropriate security services and objectives. Subsequently, proposed security solutions may be evaluated to determine if residual vulnerabilities exist, and a managed approach to mitigating risks may be proposed. *Security services* are those services that safeguard and secure information and information systems. Access Control, confidentiality, integrity, availability, and non-repudiation are the five primary security services. These services are accomplished by incorporating security mechanisms, e.g., encryption, identification, authentication, access control, security management and trust technology into the information system to form a barrier to attack. This chapter presents an overview (including a definition) of each of these services, a breakdown of the various elements included in each and a detailed look at the security mechanisms that support each service.

Three additional topics, robustness, interoperability, and KMI/PKI should be considered in selection of security countermeasures. The *robustness strategy* provides a philosophy and initial guidance for selecting the strength of security mechanisms and the security assurance provisions that may be needed for a particular value of information and a potential threat level. This section defines the IATF strategy for measuring and assessing the need for various levels of robustness for technical (and selected non-technical) security countermeasures. The robustness strategy is not intended to provide universal answers on needed strength or assurance that is, it is not a "cookbook." The final selection of mechanisms, and the necessary level of strength and assurance needed will be based on an Information Systems Security Engineering (ISSE)

activity and a resultant risk management process that addresses the situation of a specific user, mission, and environment.

The robustness of a security solution must be considered in relation to the system requirement for connectivity. Recognizing the growing need for connectivity, an *interoperability framework* provides a strategy to ensure that security provisions (1) do not inhibit the connectivity that is available without security and (2) if needed, maintain backward compatibility with existing system capabilities. The chapter continues with a discussion of *KMI/PKI Considerations*. Typically, the need for KMI/PKI capabilities accompanies the use of technical security countermeasures. It is important to consider the needs that a KMI/PKI creates and the demands it places on network users and operators in the context of any potential network security solution.

This chapter provides a basic framework for considering these important topics. Each facet of a solution addressed in this chapter should be considered in relating to the other facets. For example, the robustness of a solution depends on the way the technology is implemented. Similarly, knowledge of the primary security services and the important security technologies will facilitate the formation of effective security solutions. In addition, considering interoperability and KMI/PKI concurrently with the formulation of a security solution will help to ensure the effectiveness of that solution.

## 4.2 Adversaries, Motivations, Capabilities, and Categories of Attacks

Adversaries come from various backgrounds and have a wide range of financial resources at their disposal. In this section a host of potential adversaries are examined. What produces an adversary? What are each adversary's motivations? What category(s) of attacks does each adversary use? This section seeks to answer these questions by providing information on the various potential adversaries and by providing examples of attacks in each attack category along with a brief description of how each attack is performed and by whom.

This section also discusses the countermeasures that can be used against potential adversaries and the different categories of attack.

### 4.2.1 Potential Adversaries

One typically thinks of adversaries as having malicious intent. However, in the context of system security and protecting one's systems and information, it is also important to consider the threat posed by those without malicious intent. Table 4-1 provides examples of individuals and organizations in both of these categories.

**Table 4-1 Examples of Potential Adversaries**

Adversary	Description
<b>Malicious</b>	
Nation States	State-run, well organized and financed. Use foreign service agents to gather classified or critical information from countries viewed as hostile or as having an economic, military or a political advantage.
Hackers	A group of individuals (e.g., hackers, phreakers, crackers, trashers, and pirates) who attack networks and systems seeking to exploit the vulnerabilities in operating systems or other flaws.
Terrorists/ Cyberterrorists	Individuals or groups operating domestically or internationally who represent various terrorist or extremist groups that use violence or the threat of violence to incite fear with the intention of coercing or intimidating governments or societies into succumbing to their demands.
Organized Crime	Coordinated criminal activities including gambling, racketeering, narcotics trafficking, and many others. An organized and well-financed criminal organization.
Other Criminal Elements	Another facet of the criminal community, which is normally not very well organized or financed. Normally consists of very few individuals, or of one individual acting alone.
International Press	Organizations that gather and distribute news, at times illegally, selling their services to both print and entertainment media. Involved in gathering information on everything and anyone at any given time.
Industrial Competitors	Foreign and domestic corporations operating in a competitive market and often engaged in the illegal gathering of information from competitors or foreign governments in the form of corporate espionage.
Disgruntled Employees	Angry, dissatisfied individuals with the potential to inflict harm on the local network or system. Can represent an insider threat depending on the current state of the individual's employment and access to the system.
<b>Non-Malicious</b>	
Careless or Poorly Trained Employees	Those users who, either through lack of training, lack of concern, or lack of attentiveness pose a threat to information systems. This is another example of an insider threat or adversary.

### 4.2.1.1 Motivations

Individual motivations to “get inside” are many and varied. Those with malicious intent wishing to achieve commercial, military, or personal gain are known as “hackers”[1]. At the opposite end of the spectrum are those who accidentally do something that compromises the network. Hackers range from the inexperienced professional, college student, or novice (e.g., Script Kiddy) to the highly technical and very capable (e.g., Uberhacker). Most hackers pride themselves on their skill and seek, not to destroy, but simply to gain access so that the computer or network can be used for later experimentation. Hackers often believe that by exposing a hole or “back-door” in a computer system, they are actually helping the organization to close the holes, providing an actual benefit to the Internet and a needed resource. Other hackers have less benign motives for “getting inside”.

Intelligence gathering, information operations, or psychological warfare are some motivations behind attempts to gain access. The following are some common reasons why an adversary might be motivated to exploit a particular target.

- Gain access to classified or sensitive information. (Note: What is of high value to one person or organization may be of no value to another).
- Track or monitor the target's operations (traffic analysis).
- Disrupt the target's operations.
- Steal money, products, or services.
- Obtain free use of resources (e.g., computing resources or free use of networks).
- Embarrass the target.
- Overcome the technical challenge of defeating security mechanisms.

From an information system's standpoint, these motivations can express themselves in three basic goals: access to information, modification of or destruction of information or system processes, or denial of access to information. In attacking an information processing system an adversary accepts a certain amount of risk. This risk may be time dependent. The risk of loss to the adversary may far exceed the expected gain. Risk factors include:

- Revealing the adversary's ability to perform other types of attacks,
- Triggering responses that might prevent the success of a future attack especially when the gain is much greater,
- Incurring penalties, (e.g., fines, imprisonment, embarrassment), and
- Endangering human life.

The level of risk that an adversary is willing to accept depends on the adversary's motivation.

## 4.2.1.2 Capabilities

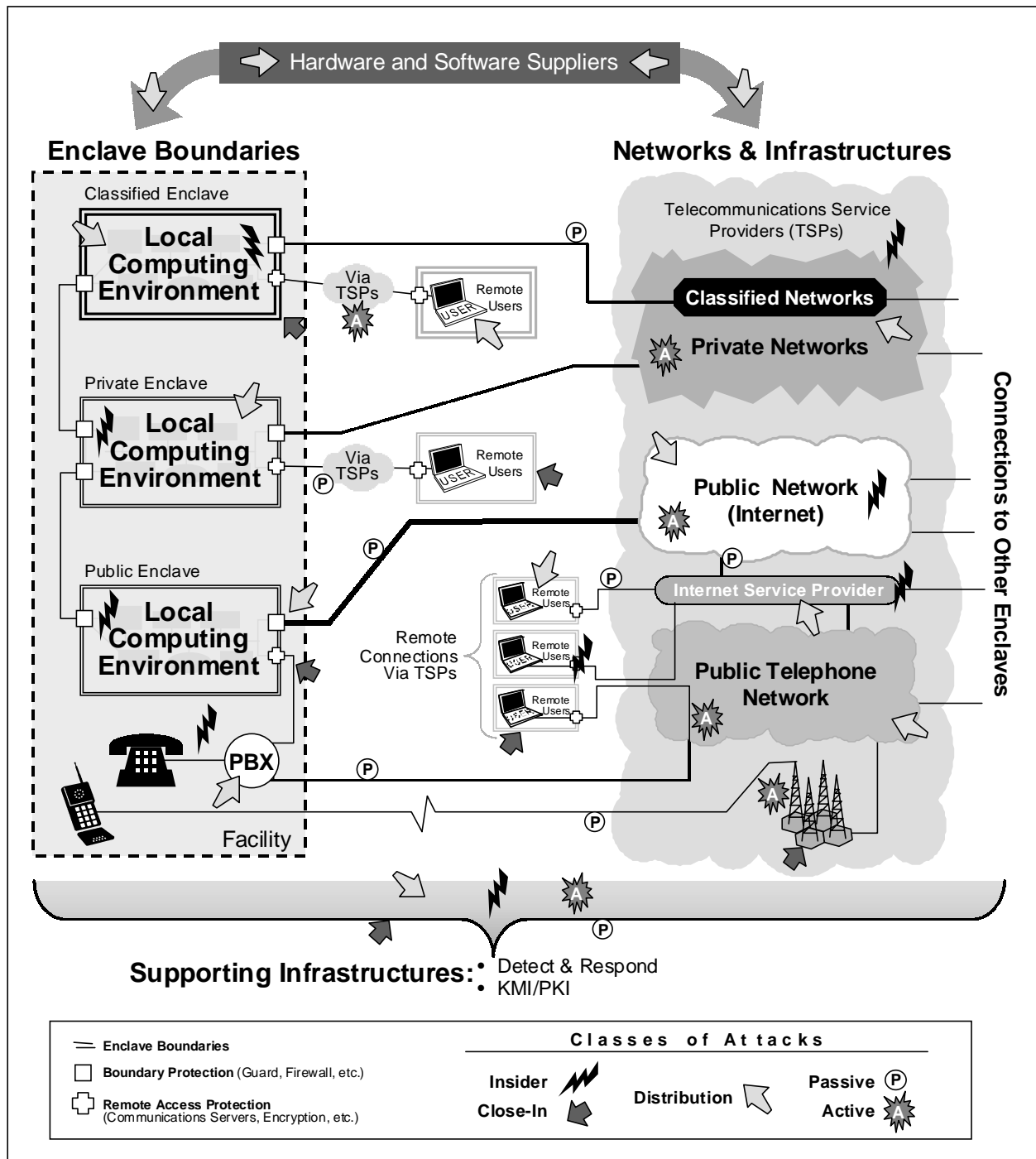
Adversaries' capabilities determine their ability to implement attacks against the information processing system. Some capability factors are:

- Knowledge and skills in developing attacks,
- Availability of necessary resources.

The greater the capabilities of the adversary, the greater the likelihood of an attack. If the adversary has the necessary knowledge, skills, and resources and is willing to put themselves and those resources at risk, then the only remaining factor is opportunity. Although opportunity is not captured in our list of capabilities it is the last key element that is necessary if an adversary is to attack. Opportunity may present itself in a number of forms including vulnerabilities in a particular operating system, misconfiguration of routers or firewalls, and modems attached but unprotected inside the local enclave. Reducing an adversary's capabilities usually is not possible but reducing the adversary's opportunity is.

## 4.2.2 Categories of Attacks

Chapter 1 (Introduction), Table 1-1, addresses five categories of system attack. As shown in Figure 4-1, each of these has unique characteristics that should be considered in defining and implementing countermeasures. This section provides an overview of each category of attack,



**Figure 4-1. Categories of Attacks Against Networked Systems**

with specific examples of attacks for each category of attack. Note that all network-based attacks are combined in the following discussion.

### 4.2.2.1 Passive Attacks

These attacks involve passive monitoring of communications sent over public media (e.g., radio, satellite, microwave, and public switched networks). Examples of countermeasures against these attacks include the use of VPNs, cryptographically protected networks, and use of protected distribution networks (e.g. physically protected/alarmed wire-line distribution network). Table 4-2 provides examples of specific attacks characteristic of this class.

**Table 4-2. Examples of Specific Passive Attacks**

<b>Attack</b>	<b>Description</b>
Monitoring Plaintext	An attacker monitoring the network could capture user or enclave data that is not otherwise protected from disclosure.
Decrypting Weakly Encrypted Traffic	Crypto-analytic capability is available in the public domain, as witnessed by the June 1997 collaborative breaking of the 56-bit-strength Data Encryption Standard (DES). While the near-term threat to large volumes of traffic is questionable given the number of machines and hours involved, it does show the vulnerability of any single transaction.
Password Sniffing	This type of attack involves use of protocol analyzers to capture passwords for unauthorized reuse.
Traffic Analysis	Observation of external traffic patterns can give critical information to adversaries even without decryption of the underlying information. For instance, extension of a network into a tactical theater of operations. Changes in traffic patterns may indicate the imminence of offensive operations thereby removing the element of surprise.

### 4.2.2.2 Active Attacks

Active attacks include attempts to circumvent or break security features, introduce malicious code (such as computer viruses), and subvert data or system integrity. Typical countermeasures include strong enclave boundary protection (e.g., firewalls and guards), access control based on authenticated identities for network management interactions, protected remote access, quality security administration, automated virus detection tools, audit, and intrusion detection. Table 4-3 provides examples of specific attacks characteristic of this class.

**Table 4-3 Examples of Active Attacks**

<b>Attack</b>	<b>Description</b>
Modification of Data in Transit	In the financial community, it could be disastrous if electronic transactions could be modified to change the amount of the transaction or redirect the transaction to another account.
Replay (Insertion of Data)	Re-insertion of previous messages could delay timely actions. Bellovin shows how the ability to splice messages together can be used to change information in transit and produce desired results.
Session Hijacking	This attack involves unauthorized use of an established communications session.

Attack	Description
Masquerade as Authorized User/Server	This attack involves an attacker identifying himself or herself as someone else, and thereby gaining unauthorized access to resources and information. An attacker gets user/administrator information by employing sniffers or other means, then uses that information to log in as an authorized user. This class of attack also includes rogue servers that can be used to obtain sensitive information after establishing what is believed to be a trusted service relationship with the unsuspecting user.
Exploit System-Application and Operating System Software	An attacker exploits vulnerabilities in software that runs with system privileges. Well-known attacks involve sendmail and X-Windows server vulnerabilities. Recently, there has been an increase in alerts regarding Windows 95 and Windows NT vulnerabilities. New vulnerabilities for various software and hardware platforms are discovered almost daily. Attacks, vulnerabilities, and patches are reported through the various computer emergency response alerts and bulletins.
Exploiting Host or Network Trust	An attacker exploits transitive trust by manipulating files that facilitate the provision of services on virtual/remote machines. Well-known attacks involve rhosts and .rlogin, which facilitate workstations sharing of files and services across an enterprise network.
Exploiting Data Execution	An attacker can get the user to execute malicious code by including the code in seemingly innocent software or e-mail that is downloaded. The malicious code might be used to destroy or modify files, especially files that contain privilege parameters or values. Well-known attacks have involved PostScript, Active-X, and MS Word macro viruses.
Inserting and Exploiting Malicious Code (Trojan Horse, Trap Door, Virus, Worm)	An attacker can gain execution access to a user's system commands through one of the vulnerabilities previously identified and use that access to accomplish the his or her objectives. This could include implanting software to be executed based on the occurrence of some future event. Hacker tools such as Rootkit (see <a href="http://www.rootshell.com">http://www.rootshell.com</a> [2] to download Rootkit or any of a large number of security/hacker tools) have turnkey capabilities, including an insertion script, root grabbing, Ethernet sniffing, and track hiding to mask the presence of a hacker.
Exploiting Protocols or Infrastructure Bugs	<p>An attacker exploits weaknesses in protocols to spoof users or reroute traffic. Well-known attacks of this type include spoofing domain name servers to gain unauthorized remote login, and bombing using Internet Control Message Protocol (ICMP) to knock a machine off the air. Other well-known attacks include source routing to impersonate a trusted host source, Transmission Control Protocol (TCP) sequence guessing to gain access, and TCP splicing to hijack a legitimate connection.</p> <p>Malicious code can exfiltrate information through a lower level tunnel within a virtual private network. At least one published paper points out potential security concerns revolving around use of Internet Protocol Security (IPSec) default security mechanisms. What Bellovin points out are occasions on which the integrity functions of DES in Cipher Block Chaining mode can be circumvented with the right applications by splicing of packets.</p>
Denial of Service	An attacker has many alternatives in this category, including ICMP bombs to effectively get a router off the network, flooding the network with garbage packets, and flooding mail hubs with junk mail.

### 4.2.2.3 Close-in Attacks

Close-in attacks are attacks in which an unauthorized individual gains close physical proximity to networks, systems, or facilities for the purpose of modifying, gathering, or denying access to, information. Gaining such proximity is accomplished through surreptitious entry, open access, or both. Table 4-4 provides examples of specific attacks characteristic of this class.

**Table 4-4 Examples of Close-In Attacks**

Attack	Description
Modification of Data/Information Gathering	This results from an individual in close proximity gaining physical access to the local system and, as a result, modifying or stealing information such as IP addresses, login ID schemes, and passwords.
System Tampering	This type of attack results from an individual in close proximity gaining access to and tampering with the system (e.g., bugging, degrading).
Physical Destruction	This type of attack results from an individual in close proximity gaining physical access, resulting in the physical destruction of a local system.

### 4.2.2.4 Insider Attacks

Insider attacks are performed by a person who either is authorized to be within the physical boundaries of the information security processing system or has direct access to the information security processing system. There are two types of insider attacks: malicious and non-malicious (carelessness or ignorance of the user). The non-malicious case is considered an attack because of the security consequences of the user's action.

- *Malicious Insider Attacks*—The Federal Bureau of Investigation (FBI) estimates indicate that 80 percent of attacks and intrusions come from within organizations (see <http://www.cs.purdue.edu/coast/intrusion-detection/>) [3]. An insider knows the layout of the system, where the valuable data is, and what security precautions are in place. Insider attacks originate from within the enclave and are often the most difficult to detect and to defend against.

Sources of insider attacks can include uncleared cleaning crews (with after-hours physical access), authorized (privileged to login) system users, and system administrators with malicious intent. Often it is difficult to prevent individuals who have legitimate access to a system from transgressing into more private areas where they have no authorization. Insider attacks may focus on compromise of data or access and can include modification of system protection measures. A malicious insider may use covert channels to signal private information outside of an otherwise protected network. However, there are many other avenues through which a malicious insider can damage an information system.

- *Non-malicious Insider Attacks*—These attacks are caused by authorized persons who have no intent to cause damage to the information or to the information processing system but may unintentionally do so by some specific action. The damage may be caused by lack of knowledge or by carelessness.



Typical countermeasures include security awareness and training; audit and intrusion detection; security policy and enforcement; specialized access control of critical data, servers, Local Area Networks (LAN), etc. implemented by trust technology in computer and network elements; or a strong Identification and Authentication (I&A) capability. Table 4-5 contains examples of specific attacks characteristic of this class.

**Table 4-5 Examples of Insider Attacks**

<b>Attack</b>	<b>Description</b>
<b>Malicious</b>	
Modification of Data or Security Mechanisms	Insiders often have access to information due to commonality of shared networks. This can give the insider access, allowing manipulation or destruction of information without authorization.
Establishment of Unauthorized Network Connections	This results when users with physical access to a classified network create an unauthorized connection to a lower classification level or sensitivity network. Typically this is in direct violation of the classified network's security policy or user directives and procedures.
Covert Channels	Covert channels are unauthorized communication paths used for transferring misappropriated information from the local enclave to a remote site.
Physical Damage/ Destruction	This is intentional damage to or destruction of a local system(s) resulting from the physical access afforded the insider.
<b>Non-Malicious</b>	
Modification of Data	This type of attack results when insiders either through lack of training, lack of concern, or lack of attentiveness, modify or destroy information located on the system.
Physical Damage/ Destruction	This type of attack is listed under malicious as well. As a non-malicious attack, it can result from carelessness on the part of the insider, for instance, failure to obey posted guidance and regulations resulting in accidental damage to or destruction of a system.

## 4.2.2.5 Distribution Attacks

The term “distribution attack” refers to the potential for malicious modification of hardware or software between the time of its production by a developer and its installation or when it is in transit from one site to another. The threat at the factory can be minimized by strong in-process configuration control. Distribution threats can be addressed by use of controlled distribution, or by signed software and access control that is verified at the final user site. Table 4-6 contains examples of specific attacks characteristic of this class.

**Table 4-6 Examples of Distribution Attacks**

Attack	Description
Modification of Software/Hardware at Manufacturer's Facility	These attacks can be performed by modifying the configuration of software or hardware while it is cycling through the production process. The countermeasures for threats during this phase include rigid integrity controls including high-assurance configuration control, and cryptographic signatures on tested software products.
Modification of Software/Hardware during Distribution	These attacks can be performed by modifying the configuration of software or hardware during its distribution (e.g. embedment of listening devices during shipment). The countermeasures for threats during this phase include use of tamper detection technologies during packaging, use of authorized couriers and approved carriers, and use of blind-buy techniques.

## 4.3 Primary Security Services

The IATF guidance incorporates five primary security services: access control, confidentiality, integrity, availability, and non-repudiation. The division of network security principles into standard security service categories is convenient for this description. The categories presented below roughly follow the “basic security services” identified in the 1990 Recommendation X.800, “Security Architecture for Open Systems Interconnection for Consultative Committee for International Telephone and Telegraph (CCITT) Applications” (which is technically aligned with ISO 7498-2, “Information Processing Systems Open Systems Interconnection, Basic Reference Model”, Part 2: Security Architecture) and more recently, the International Organization for Standardization (ISO)/International Engineering Consortium (IEC) 10181 series, Parts 1-7.

In practice, not one of these security services is isolated or independent of the other services. Each service interacts with and depends on the others. For example, confidentiality is of limited value unless preceded by some type of authorization process. Protection from unauthorized entities is not possible if one cannot determine whether the entity he or she is communicating with is authorized. In actual implementations, lines between the security services are blurred by the use of specific mechanisms that contribute to supporting more than one service.

Given this caveat, this section characterizes each service according to its basic functional elements and discusses the mechanisms that are available to implement the elements of that service. Where appropriate, considerations of the relative strengths of these mechanisms are also noted.

### 4.3.1 Access Control

In the context of network security, access control means limiting access to networked resources (hardware and software) and data (stored and communicated). The goal of access control is to prevent the unauthorized use of these resources and the unauthorized disclosure or modification of data. Access control also includes “resource control”, for example, preventing logon to local workstation equipment or limiting use of dial-in modems. For the purposes of this discussion, network access control is not concerned with denying physical access (e.g., via locked rooms or tamperproof equipment).

Access control is applied to an “entity” based on an Identity and/or an Authorization. An identity may represent an actual user, a process with its own identity (e.g., a program making a remote access connection), or a number of users represented by single identity (e.g., role based access control).

Access control mechanisms are most often used as a set of mechanisms, which may be used by other security services. Confidentiality, integrity, availability, and limiting use of network resources all depend on limiting the ability of an adversary to access an item or service.

The elements of access control can be categorized in the following areas.

- *I&A* – Establishing the identities of entities with some level of assurance (an authenticated identity).
- *Authorization* – Determining the access rights of an entity, also with some level of assurance.
- *Decision* – Comparing the rights (authorization) of an authenticated identity with the characteristics of a requested action to determine whether the request should be granted.
- *Enforcement* – Enforcement may involve a single decision to grant or deny or may entail periodic or continuous enforcement functions (continuous authentication).

The following subsections discuss these elements and provide examples of the mechanisms that are available to implement them.

### 4.3.1.1 I&A

I&A is a set of security services used in conjunction with most other security services. The first step of most security services is to determine the identities of one or more of the parties participating in an action. A trusted identity must be used for access control decisions as well as to provide non-repudiation and accountability evidence. Knowing the identity of an entity and the existence of a peer relationship is also fundamental to establishing communication with confidentiality and integrity. If the identity of the peer in a secure communications path is not properly established, it leaves open the possibility that an unauthorized user (an adversary) could masquerade as an authorized user, leaving the data open to disclosure or manipulation by the adversary.

The process of determining an authentic identity consists of the following.

#### 4.3.1.1.1 Assigning, Binding, and Representing

There must be a mechanism for providing some assurance in the assignment of an identity. The entity that assigns identity must have a position with some level of trust (either implied or assured by a third entity common to both with a higher position or level of trust. These trusted entities must implement a process of identity checking that protects against assignment of improper identities. Process examples include checking driver’s licenses or verifying fingerprints. Assigning identity is the equivalent of a registration process and could take place through an existing security mechanism with its own identity establishment mechanism.

An identity must be unique amongst the community that will be validating that identity. This requires implementation of a community wide assignment mechanism that provides a unique identity to each entity. An identity also must be conveyed in a representation that is recognized and can be commonly processed by all potential peers to a communication. This implies a standard format for representing identity.

Identities used for network access control can be assigned and represented by many different mechanisms; for example:

- System administrators providing accounts and passwords for UNIX usernames,
- Network administrators assigning Internet Protocol (IP) addresses to machines,
- Key distribution methods that distribute symmetric keys,
- Key distribution methods that distribute public/private key pairs,
- Certification Authorities (CA) generating public key certificates containing Distinguished Names (DN), and
- Security officers associating a set of fingerprints with a common name.

The assurance level attributed to an identity depends on the processes used to verify the correctness of that identity before it is issued, the trust instilled by the entity assigning the identity, and the strength of the binding between the entity and the identity. Verification may range from requesting a mother's maiden name over the telephone to checking driver's licenses or verifying fingerprints in person. Means of instilling trust in issuers include procedural mechanisms, such as a company's assigning system administrators, legal mechanisms, such as notaries, and technological mechanisms such as certification paths in a certification hierarchy. Mechanisms for binding entities to identities include signed X.509 certificates and password files associated with access control lists.

Strongly establishing identities for communicating entities is the first step in countering any attack that is predicated on adversaries, representing themselves as someone or something that they are not (including masquerading and insider modification attacks).

### 4.3.1.1.2 Communicating and Authenticating

To authenticate an entity that is attempting to gain access, an identity must be associated with the access request and provided to the communicating peer. Along with an indication of identity, the authenticating peer must have the parameters (authentication information) needed to validate that identity. Authentication is implemented by user-to-host and peer-to-peer, and Trusted Third Party (TTP) architectures as follows.

- User-to-Host—When a user logs onto a host (or workstation), the user must be identified and authenticated before access to the host or network is granted. This process requires a mechanism to authenticate a real person to a machine. The best methods of doing this involve multiple forms of authentication such as password, physical token, and biometric verification (e.g. something you *know*, something you *have*, something you *are*).
- Peer-to-Peer Authentication—A peer-to-peer authentication architecture, sometimes referred to as mutual authentication protocol, involves the direct communication of

authentication information between the communicating entities (e.g., peer-to-peer or client host-to-server). No other entities are required. This architecture is possible only if each entity in a security domain is able to obtain the authentication information of every communicating entity in the domain.

- **Trusted Third Party Authentication**—The architecture for TTP authentication uses a third entity, trusted by all entities, to provide authentication information. A TTP may provide authentication information in each instance of authentication, in real-time, or as a precursor to an exchange (such as a certificate authority). The amount of trust given the third party must be evaluated. Methods of establishing and maintaining a level of trust in a TTP include certification practice statements that establish rules, processes, and procedures that a CA uses to assure the integrity of the authentication process and use of secure protocols to interface with authentication servers.

The mechanisms used for authenticating of an identity can be categorized as simple or cryptographically based. Simple mechanisms may include identification based on identities which are verified by asking the entity to communicate information that only the entity attempting access would know (e.g. a password and locally stored password file). Assurance comes from the local binding between the password and an identity. Another example of a simple authentication method is address-based authentication. Address-based mechanisms authenticate identity based solely on assigned network addresses (e.g., Internet Protocol (IP) address) of communicating peers as compared to known IP address assignment for the entities.

Cryptographic-based mechanisms rely on the cryptographic processing of data within a defined protocol. Peers may share a common secret key (often stored in a hardware token) to process, or encrypt the exchange in a challenge-response protocol. Other cryptographic mechanisms rely on public key cryptography alone, or on the binding between a public key and an identity provided by public key certificates. Examples of how an identity is authenticated in each cryptographic technique are provided below.

- **Identity Is A Locally Defined Name**—Identities of all potential communicating peers are stored locally in a trusted database that associates identities with their public keys. These public keys correspond to the private key used to sign a unique piece of data. Verifying a signature by using a stored public key authenticates an identity.
- **Identity Means the Defined Name**—From the valid X.509 certificate containing the public key that corresponds to the private key used to sign a unique piece of data. A valid X.509 certificate means that the complete certification path has been validated (including Certificate Revocation List (CRL) and Compromised Key List (CKL) checks and validity periods for all certificates) to a trusted root. X.509 certificates (of communicating peers or of the entities in certification paths) may be stored locally (cached), carried in the security association protocol, or accessed as needed from an X.500 directory, or any combination of these three methods. Verifying a signature by using a valid public key authenticates an identity.

For all cryptographically based mechanisms, the strength of the mechanism lies partly in the strength of the cryptographic algorithms (including key size), partly in the security of any communications protocol, and in large part, in the protection provided to secret key material.

There are a number of mechanisms for implementing and distributing identity and authentication information. Some of these mechanisms are:

- Names and passwords stored in a database local to the entity making the access control decision.
- IP addresses provided by a secure Domain Name Server (DNS).
- Passwords generated locally based on time (one time passwords).
- Symmetric keys stored in a local database.
- Public keys stored in a local database,
- Public key certificates provided by directories in response to queries.
- Authentication information may be carried in the communications protocols themselves.

The assurance level of the communication of identity and authentication information processes depends on whether that information needs protecting and how well it is protected. For example, passwords are sensitive because they can be used by anyone who knows them; they should therefore be encrypted for storage and transport. Certificates can be stored in unprotected directories or carried on unencrypted communications channels because they can only be used by the entity that holds the associated private key.

Note that identity information and the information used to authenticate that identity do not have to flow over the same communications path. A common example is name and password logins. Users are queried for a name and an associated password (the identity information) over the communications protocol. The authenticity of that name and password pair is established only by checking a locally stored database (the information used to authenticate provided by an off-line process).

There are entire infrastructures devoted to providing identities and the means of authenticating those identities. Examples of infrastructures supporting the determination of an authentic identity include the X.509 authentication framework, the Internet Engineering Task Force (IETF) public key infrastructure, the secure DNS initiatives, and the Simple Public Key Infrastructure (SPKI).

### 4.3.1.2 Authorization

Another important step in an access decision is determining the authorizations of one or more of the parties participating in a communication. These authorizations result in the granting of a set of privileges to an entity. Much like identity, authorizations must be conveyed in a commonly understood format and must be presented or maintained with some level of confidence. The process of determining an authenticated set of authorizations generally consist of the same components as determining an authenticated identity. A strong mechanism for determining authorizations can prevent an attack in which an entity attempts to forge access rights from being successful.

The process of determining the authorizations of an entity consists of assigning authorizations, binding authorizations to an entity, representing those authorizations in a standard format,

communicating those authorizations, and establishing the authenticity of the authorizations. These steps are discussed below.

#### 4.3.1.2.1 Assigning, Binding, and Representing

As in assigning identity, the process that determines and assigns authorizations must evoke a level of trust. Responsibility for that process falls on roles with names like certificate authority (CA), attribute authority, Access Control List (ACL) administrator, and system administrator. Authorizations used for network access control can be assigned by:

- System administrators who assign usernames to groups,
- Data owners who grant authorizations to read/write/execute files,
- Network administrators who generate ACLs,
- X.500 CAs who generate version 3 X.509 certificates containing extensions, and
- Attribute authorities who generate ANSI X9.57 attribute certificates.

#### 4.3.1.2.2 Communicating and Authenticating

Communicating authorization information follows the same model as authentication information. It may be pre-distributed and stored at each entity (e.g., ACLs); it may be carried in the communications protocol; or it may be provided by a trusted third party (e.g., X.500 directory, Radius authentication servers). There are a number of models for distributing authorization information:

- ACLs stored local to the entity making the access control decision.
- X.500 directories deployed to provide X.509 certificates.
- X.500 directories deployed to provide attribute certificates.
- Authenticity of authorization information is provided either by its trusted relationship with identity information (local binding) or because it is carried in cryptographically verifiable certificates.

The level of trust attributed to the third parties used for obtaining authorization information (either the parties who generated authorizations initially or those that distribute them when needed) is always an issue. The cryptographic techniques invoked to prove the authenticity of X.509 certificates and to bind attribute certificates to identity certificates represent one attempt to assure that trust.

#### 4.3.1.3 Access Control Decision

The components discussed previously provide the information required to make an access control decision. They provide mechanisms for determining both the identity and the privilege set of a communicating entity. In practice, access decisions are usually based on an access control policy commonly referred to in the classified arena as discretionary or mandatory policies. International standards do not use the "mandatory/discretionary" terminology but rather Identity Based Access Control (IBAC), which bases decisions on an identity or Rule Based Access Control (RBAC), which checks an entity's authorizations against an established rule set.

Within the scope of this discussion, IBAC and discretionary policies can be considered equivalent, and RBAC and mandatory policies can be considered equivalent. In either case, the function of an access control decision is to result in a simple grant or deny response to requests for access.

An IBAC decision grants or denies a request based on the presence of an entity on an ACL. If an entity is on the ACL, access to the requested information or resource is permitted—otherwise, access is denied. IBAC requires an authenticated identity prior to granting any access.

An RBAC decision depends on policies that can be algorithmically expressed and thus implemented on a computer system. These policies are stated such that resources have restrictions and entities must have authorizations. Access is granted to a resource based on an entity's authorizations rather than an entity's identity. An RBAC decision requires authorization information and restriction information to compare before any access is granted.

A composite policy, referred to as role-based policy, can be considered a variation of both IBAC and RBAC. An identity is assigned to a group that has been granted authorizations. Identities can be members of one or more groups. A current example is the Global Command and Control System (GCCS), which depends on organizational and role associations.

Most network operating systems have their own method of implementing access control, but they are all identity-based IBAC. Entities are granted access to resources based on an identity established during network login, which is compared with one or more ACL. These lists may be individually administered, may be centrally administered and distributed to individual locations, or may reside on a central server(s).

Mechanisms for establishing identities and authorizations are discussed in previous sections. Mechanisms for establishing the restrictions on a resource must be provided to implement an RBAC scheme. Since rule-based access controls how rules are implemented primarily in systems dealing with sensitive information, restrictions are most often expressed as policies for accessing sensitive data. To facilitate these policies, the sensitivities of a data item are conveyed in a data label and must be compared with the set of privileges assigned to an entity. Access is granted to sensitive information if an entity's privileges are appropriate for the sensitivities of the data. An example of a rule-based policy is the classifications used to distinguish information on a national security level, such as Top Secret, Secret, and Confidential, and the rule that identities granted authorization for any security level are authorized access to all lower security levels. Users who hold Secret clearances will be allowed to access Secret and below classified information.

Consistent with the issues surrounding identities and authorizations, data labels must also be assigned, bound, represented, communicated, and authenticated. There are currently many representations of a data security label (Federal Information Publications (FIPS)[4] 188 Standard Security Label, SDE Security Label—IEEE 802.10g, Internet Security Label, International Organization of Standardization (ISO) SC-27 Security Label, Common Security Label (Military Standard [MIL STD] 2045-48501), X.411 MHS: MTS Service Definition—Security Label). Establishment of a universally accepted standard is an area for further work.

Note that an access request can actually be composed of a complicated set of parameters. For example, a particular access might be- Execute a file labeled Top Secret at 3:15 p.m. during a



time of war. Defining “access” in this manner allows the access decision function to provide a binary “grant” or “deny” result. This introduces a new set of information that must be represented, communicated, and authenticated including contextual information, such as time, status, or current conditions.

### 4.3.1.4 Enforcement

Actual enforcement of the access control decision is really the step that provides the protection against threats. All previously discussed mechanisms for preventing attacks come together here with the enforcement of those protections.

The concept of enforcing an access control decision is separate from the decision itself. This is because the two processes may reside in different places architecturally. This separation permits the concept of an “authentication server” that makes an access decision for the network communications process to allow or prevent a requested access from taking place. For example, the access decision may result in the subject’s being provided with a token (such as a certificate) that guarantees the subject the right to access its target up to but no more than  $n$  times before a given time. This token is called a ticket or capability. These tokens may be cached at the target to improve efficiency.

An access control decision and its enforcement can be made at either end of a communications association. An example is the difference between a client accessing a File Transfer Protocol (FTP) server (the server limits access to files after a client request is submitted) and an electronic mail (e-mail) message (in which the originator decides whether the recipient should receive the message before a connection is made). In the e-mail example, the recipient’s mail software may also perform an additional access control check to determine whether the recipient can be allowed to view the message.

Another distinction between access control mechanisms is whether the decision and enforcement process occurs once at the initiation of a communications session, is repeated periodically throughout a session, or qualifies as “continuously authenticated.” A method commonly used to assure that access to a communications session is controlled continuously is use of encryption mechanisms to prevent loss of control of the session (session stealing or hijacking). Indeed, it can be argued that access is not completely controlled if information flowing over a public network is not protected by the confidentiality security service.

Enforcement of an access control decision may take place at many places in a network’s architecture. Access controls may be enforced at network boundaries (e.g., firewalls, routers, and dial-in communications servers) at application servers, or anyplace in the protocol stack or operating system of individual workstations. An important implementation option is to include access control mechanisms at many layers throughout a network architecture.

## 4.3.2 Confidentiality

The confidentiality security service is defined as preventing unauthorized disclosure of data (both stored and communicated). This definition is similar to (and actually a subset of) the description of access control in Section 4.3.1. In fact it can be argued that providing access control also provides confidentiality or conversely, that providing confidentiality is a type of access

control. We include in the definition of “information”, data that is not traditional user data—(examples are network management data, routing tables, password files, and IP addresses on data packets). Confidentiality services will prevent disclosure of data in storage, transiting a local network, or flowing over a public Internet. One subset of confidentiality is that of anonymity, a service that prevents disclosure of information that leads to the identification of the end user.

The requirements for the provision of the confidentiality security service depend on a number of variables:

- *The Location(s) of the Data that Needs Protection*—Data can exist on, an individual machine (e.g., on a hard disk in an end system or in a file on a server) on the wires of a local network; in transport via other mechanisms (e.g., floppy disk) or flowing across a totally public medium (e.g., across the Internet or via a satellite).
- *The Type of Data that Needs Protection*—Data elements may be local files (e.g., passwords, or secret keys) data carried in a network protocol, or the exchanges of a network protocol, (e.g., a protocol data unit).
- *The Different Amounts or Parts of User Data that Need Protection*—It may be necessary to protect an entire data element or only parts of a data element or protocol data unit or the existence of an entire set of protocol exchanges.
- *The Value of the Data that Needs Protection*—The sensitivity and perishability of the data being protected influence the provision of security services, particularly the strength of mechanisms implemented. The value of the data both to the owner and to potential threats must be considered.

The elements of confidentiality are as follows.

- *Data Protection*—This is prevention of disclosure of the contents of data even if it is accessible (e.g., flowing over a network). This element invokes mechanisms that act directly on the data (or act in response to characteristics of the data) rather than acting in response to an entity’s attempt to access data.
- *Data Separation*—Data separation traditionally refers to the concept for providing for separate paths (red/black or physical) or process separation (COMPUSEC techniques, etc).
- *Traffic Flow Protection*—Data characteristics include frequency, quantity, destination of traffic flow, etc. Traffic flow protection includes not only characteristics but inference information such as command structure, and even the instance of communication (e.g., a network communication).

### 4.3.2.1 Data Protection

In cases in which communicated data will be visible to possible adversaries (i.e., via passive monitoring attacks), the most common method for providing confidentiality via data protection is to encrypt the appropriate data. Encryption is also used to protect stored data that might be accessed by an adversary (e.g., via the network-based attacks described in Chapter 3 (Information Systems Security Methodology)).

Encryption is defined as the transformation of data into a form that is unreadable by anyone who does not possess the appropriate secret key. There are many examples of using encryption to provide confidentiality. A small subset includes:

- Security-enabled applications (file encryptors),
- Secure peripherals (media encryptors),
- Operating systems (encrypt local passwords),
- Secure application protocols (secure File Transfer Protocol (FTP)),
- Security protocols (authentication and key management protocols),
- Secure upper layer network protocols (socket layer, IP layer), and
- Secure lower layer network protocols (link encryptors).

There are two types of cryptographic mechanisms that can be used to provide an encryption capability: symmetric cryptography wherein entities which share a common secret key, and public key cryptography (also known as asymmetric cryptography) in which each communicating entity has a unique key pair (a public key and a private key).

Implementation variables in providing encryption for protection of communications data include where in the protocol stack encryption takes place. Encryption at different layers provides different protections to the underlying data or protocol elements.

The strength of the confidentiality service may depend on a number of variables associated with the encryption function.

- The security protocol or application used to invoke the encryption function.
- The trust in the platform executing the protocol or application.
- The cryptographic algorithm.
- The length of the key(s) used for encryption/decryption.
- The protocol used to manage/generate those keys.
- The storage of secret keys (key management keys and encryption keys).

### 4.3.2.2 Data Separation

Data separation takes a different approach to preventing disclosure. Mechanisms that provide data separation prevent the adversary from ever getting at the data in the first place. This is achieved using the normal access control mechanisms described in Section 4.4 (Important Security Technologies), as well as by the additional techniques described below. An example of a commonly used data separation technique is to not allow data labeled as Secret to flow onto an unclassified network.

Data separation mechanisms provide confidentiality by preventing data from reaching a location or destination where it could be disclosed to unauthorized entities. Mechanisms can be employed to prevent data from flowing into undesired areas (routing control). Other mechanisms may be employed to physically segregate those areas. Examples of routing control include a router that directs IP packets based on security labels, thereby preventing Secret packets from reaching unclassified networks, and a firewall that scans e-mail messages for

“dirty words” and prevents messages containing them from being released outside a local network. Examples of physically segregated data are isolated system high networks and physically protected wires.

Data separation mechanisms can be used to counter passive monitoring attacks, as well as insider attacks that inappropriately attempt to release information from a controlled area. The primary variable in the level of assurance provided by a data separation mechanism is the level of trust associated with the process or machine implementing the mechanism.

### 4.3.2.3 Traffic Flow Protection

Data padding can be employed to provide traffic flow protection. Addition of superfluous (usually random) data to data carried in a communications protocol can hide the characteristics (e.g. data rate, data frequency, etc.) of that underlying data. When combined with encryption, this mechanism also hides the content of the underlying data.

Address hiding may also be employed to provide traffic flow protection. Address hiding includes network address translation in which the IP addresses of machines in a local network are replaced by the address of a protecting firewall. Network layer addresses may be hidden by encrypted tunnels, which also provide data confidentiality.

### 4.3.2.4 Other Mechanisms

Other mechanisms for providing confidentiality include spread-spectrum and frequency hopping techniques.

## 4.3.3 Integrity

The integrity security service includes one or more of the following: prevention of unauthorized modification of data (both stored and communicated), detection and notification of unauthorized modification of data, and the journaling of all changes to the data. Modifications to both stored and communicated data may include changes, insertions, deletions, or duplications. Additional potential modifications that may result when data is exposed to communications channels include sequence changes and replay.

The requirements for the provision of the integrity security services are similar to those of confidentiality. These include the location, type, and amount or parts of the data that needs protection.

When integrity is discussed with respect to network security, it is important to consider where in the protocol stack the integrity service is provided. Different implementation (layering) options will provide integrity to different protocol layer data as well as the data being communicated. Sophisticated integrity schemes are likely to require service from the application utilizing the data.

Note that integrity protection is of no value unless it is combined with a mechanism that provides authentication of the source. Without source authentication, anyone could have tampered with the original data and then just reapplied an integrity mechanism.

Data integrity can be divided into two types, based on the type of data to be protected. Integrity can be applied to a single data unit (protocol data unit, database element, file, etc.) or to a stream of data units (e.g., all protocol data units exchanged in a connection).

### 4.3.3.1 Single Unit of Data

Assuring the integrity of a single data unit requires that the originating (sending) entity calculate an additional data item that is a function of (and bound to) the original data unit. This additional item is then carried along with the data unit. The entity that desires to verify the integrity of this data unit must recalculate the corresponding quantity and compare it with the transferred value. A failure of the two to match indicates that the data unit has been modified in transit.

Methods for calculating this piece of data which is a function of the original data unit (the “check value”), vary in the processing required and services provided. Checksums, Cyclic Redundancy Check (CRC) values, and hashes (also known as a message digest) all meet the requirement that they depend on the entire content of the original data unit. A weakness of this method is that, if an adversary modifies the original data, these functions are easily reproducible and allow the adversary to generate a proper value (for the modified data) thereby defeating the integrity service. An additional mechanism can be applied to prevent access to the check value (e.g., encryption or digital signatures) to overcome this problem.

Another method of preventing (successful) modification of the check value is to include a secret value along with the original data unit. This property is exhibited by Message Authentication Codes (also known as Message Integrity Check and keyed hashes).

The icheck value alone will not protect against an attack that replays a single data unit. A time stamp may be included along with the original data unit to provide limited protection against replay.

### 4.3.3.2 Sequence of Data Units

To protect the integrity of a sequence of data units (i.e., protect against reordering, losing, replaying and inserting, or modifying data), some type of ordering information must be provided within the communications protocol. Examples of ordering information are sequence numbers or time stamps. Integrity of sequences can also be provided by encrypting the sequence of data units using a cryptographic algorithm in which encryption of each sequence depends on the encryption of all previous sequences (also referred to as chaining).

## 4.3.4 Availability

Availability is “the property of being accessible and usable upon demand by an authorized entity.” Availability in a networked environment includes not only the user’s ability to access hardware and software resources (such as user agents and servers) but also the user’s ability to obtain a desired Quality of Service (QoS); (e.g., make use of network bandwidth with reasonable throughput). Network traffic must be able to traverse local and wide area networks as required to reach its intended destination.

One of the most effective methods of assuring availability is to provide a secure network environment that exhibits the common security services. Attacks that could prevent a networked system from providing availability may be countered by preventing unauthorized access to resources with access controls and protecting data from disclosure or modification with integrity and confidentiality services. Access control, integrity, and confidentiality become mechanisms to help support the availability security service.

Solutions to problems that affect availability include the following.

- *Protection from Attack*—Some network-based attacks are designed to destroy, degrade, or “crash” network resources. The solution is to harden these resources against such attacks. Means of doing this include closing security holes in operating systems or network configurations, limiting access to resources to authorized entities, and limiting an adversary’s ability to manipulate or view the data flowing through and to those resources (thus preventing insertion of harmful data, such as viruses, or disclosure of sensitive network data, such as routing tables).
- *Protection from Unauthorized Use*—Availability is also limited if a resource is in use, occupied, or overloaded. If unauthorized users are using limited resources (e.g., processing power, network bandwidth, or modem connections), the resources are not available for authorized users. Identifying and authenticating the users of these resources can provide access controls to limit unauthorized use. However, the process of requesting IA to frequently may be used to slow or stop network operations (i.e., non-delivery notice floods).
- *Resistance to Routine Failures*—Normal operational failures and acts of nature also contribute to loss of availability. Solutions include use of equipment designed for high reliability, redundancy in equipment, and network connectivity that provides multiple routes.

Trusted operating system concepts are also used to limit the harmful effects of network attacks. By containing the damage done by malicious code and assuring the proper operation of other security mechanisms, the trusted operating system preserves availability. Another feature exhibited by trusted operating systems is process integrity. This provides assurance that processes executing on an end system provide consistent, repeatable results that are not affected by undesired (unauthorized) influences.

Critical system components must also provide physical security, not only to prevent attacks or misuse of resources, but also to ensure that the platforms and applications are not modified to bypass the invocation of those security services that provide availability.

### 4.3.5 Non-Repudiation

Repudiation is denial by one of the entities involved in a communication that it participated in that communication. The non-repudiation security service provides the ability to prove to a third party that the entity did indeed participate in the communication. When discussed in the context of networking.

- *Non-repudiation with proof of origin* provides the recipient of a data item with proof of the identity of the originator of that data item and the time of origination.

- *Non-repudiation with proof of delivery* provides the originator of a data item with proof that the data item was delivered to the intended recipient (and in some cases, the time of receipt).
- *Auditing services* help provide accountability of the parties involved in exchanges requiring non-repudiation by recording relevant events that can be traceable to persons that can be held responsible for their actions.

The non-repudiation service is primarily provided by application layer protocols. Users are most often concerned with providing non-repudiation for application data (such as an e-mail message or a file). Providing non-repudiation at a lower protocol layer will only provide proof that a particular connection was made; it will not bind the data that flowed over that connection to a particular entity.

Non-repudiation is provided by the authenticating characteristics of digital signatures. A digital signature on a data element (or on the hash of that element) irrevocably ties that data element to the identity contained in the public key certificate associated with the private key that generated the signature. Of course, data integrity must be provided to that data element to assure that the element was not changed after the application of the signature.

Since non-repudiation depends on an identity contained in a public key certificate (and certificates become invalid), it is important to be able to establish, to a third party, the validity of the certificate. It must be possible to prove the validity of that certificate at the time of the original communication as well as any time in the future. This can be accomplished with a combination of trusted time stamps, third party notaries, or archived certificate revocation lists.

Time-stamping achieves the goal of establishing the time at which a communication or transaction occurred. For the highest levels of assurance, time-stamps are applied by a trusted time-stamping service that digitally signs the data item (or a hash of the data item) along with the time stamp before delivery to the intended recipient.

## 4.4 Important Security Technologies

An overview of technical security countermeasures would not be complete without at least a high-level description of the widely used technologies underlying those countermeasures. This section highlights selected technologies as an introduction to the detailed technology assessments included in Chapters 5 through 9 (Defend the Network and Infrastructure, Defend the Enclave Boundary/External Connections, Defend the Computing Environment, Supporting Infrastructures, and Information Assurance for the Tactical Environment). For convenience, these technologies are listed alphabetically.

- **Application Layer Guard**—The need for a separate mechanism to perform a gatekeeper function, checking the invocation of security features, gives rise to a need for security at the application layer. This gatekeeper has recently taken the form of an application layer guard that implements firewall mechanisms (performing I&A functions and enforcing security policies such as allowing or disallowing connections based on identification and/or requested protocol processing). Guard functionality includes such features as performing a cryptographic invocation check on information that is allowed outside the protected enclave, and data content filtering to support sensitivity regrade

decisions. The guard functionality, while effective for non-real-time applications (e.g., e-mail) on networks with low sensitivity, has been difficult to scale to highly classified networks and real-time applications.

- **Application Program Interface (API)**—APIs are a means of isolating a computing platform from the details of the implementation of cryptographic functions (both the actual algorithms and the hardware implementations). It provides standard interfaces so that multiple vendors may provide interoperable solutions.
- **Common Data Security Architecture (CDSA)**—The CDSA is a set of layered security services that address communications and data security problems in the emerging Internet and intranet application space. CDSA focuses on security in peer-to-peer distributed systems with homogeneous and heterogeneous platform environments. The architecture also applies to the components of a client/server application. The CDSA addresses security issues and requirements in a broad range of applications by:
  - Providing layered security mechanisms (not policies);
  - Supporting application-specific policies by providing an extensibility mechanism that manages add-in (policy-specific) modules;
  - Supporting distinct user markets and product needs by providing a dynamically extensible security framework that securely adds new categories of security service;
  - Exposing flexible service provider interfaces that can accommodate a broad range of formats and protocols for certificates, cryptographic keys, policies, and documents; and
  - Supporting existing, secure protocols, such as Secure Sockets Layer (SSL), Secure/Multipurpose Internet Mail Extension (S/MIME), and Secure Electronic Transaction (SET).
- **Circuit Proxy**—Circuit gateways are another type of proxy firewall. A circuit-level proxy becomes an intermediate connection point in a session between a client and a server. To reach a distant server, a client initially connects to a TCP port on the circuit proxy machine. The circuit proxy then completes the connection (after making an access control decision) to the target server. Access controls are based on the identity of the initiating machine without interpreting the application protocol or viewing the contents of protocol packets. A circuit-level proxy can be utilized across several application protocols; however, client modifications may be necessary to use the circuit-level protocol.
- **CryptoAPI**—The Microsoft Cryptographic API provides services that enable application developers to add cryptography to their Win32 applications. Applications can use the functions in CryptoAPI without knowing anything about the underlying implementation, in much the same way that an application can use a graphics library without knowing anything about the particular graphics hardware configuration.
- **Cryptographic Service Providers (CSP)** —Both CDSA and CryptoAPI make use of the concept of CSPs which are independent modules that perform the real cryptographic work. Ideally, CSPs are written to be completely independent of any particular application, so that a given application will run with a variety of CSPs. In reality, however, some applications may have very specific needs that require a custom CSP.



A CSP may implement one or more of the following cryptographic functions: bulk encryption algorithm, digital signature algorithm, cryptographic hash algorithm, unique identification number, random number generator, secure key storage, and custom facilities unique to the CSP.

A CSP may be implemented in software, hardware, or both. A CSP or an independent module can also deliver key management services, such as key escrow or key recovery. CSPs should not reveal key material unless it has been wrapped. Also, the key-generation function of a CSP should be made as tamper-resistant as possible.

- **File Encryptors**—These provide confidentiality and integrity for individual files, provide a means of authenticating of a file's source, and allow the exchange of encrypted files between computers. File encryptors typically implement a graphical user interface (GUI) that allows users to choose files to be encrypted or decrypted. This protects individual files, but it does not protect all of the files on the drive.

Many applications generate temporary files that may contain user data. These files are normally erased when the application is closed; but when the application does not close in an orderly fashion, these temporary files may remain. Some operating systems do not actually erase data when files are deleted. Instead, they alter the name of the file in the file allocation table. The user's data remains on the hard drive until the space is reallocated to another file and overwritten. Thus, unencrypted and potentially classified user data can remain on the hard drive after system shutdown, either by failure to erase temporary files or by design of the operating system's erasing function.

- **Hardware Tokens**—A number of hardware token approaches are available ranging from an external memory device to one with significant levels of processing. One hardware token that is prominent in the Department of Defense (DoD) community is the FORTEZZA® Crypto Card. The FORTEZZA® card provides the cryptographic algorithms required to provide security services to a FORTEZZA® based system. It stores the private key information for each user personality, the certificates of its issuers, and the public keys needed for cryptography. It performs the digital signature and hash algorithms, public or private key exchange functions, encryption, and decryption. The interface to the card depends on the hardware platform and its configuration, and the operating system.
- **Intrusion and Penetration Detection**—Intrusion detection and response systems can protect either a network or individual client platforms. Effective intrusion systems detect both insider and outsider threats. In general, intrusion systems are intended to protect against and respond to situations in which the available countermeasures have been penetrated, either through allowed usage or the exploitation of vulnerabilities that are unknown or have not been patched. The objective of these systems is to detect malicious and unintended data and actions (e.g., altered data, evil executables, requests that permit unintended resource access, and unintended use of intended services). Once the intrusion is detected, an appropriate response is initiated (e.g., disconnect attacker, notify operator; respond automatically to halt or lessen the attack; trace attack to proper source; and counter the attack, if appropriate). Intrusion detection mechanisms operating at the transport layer can view the contents of transport packets (e.g., TCP packets) and are able to detect more sophisticated attacks than mechanisms that operate at the network layer. Intrusion detection mechanisms operating at the network layer can view

the contents of network packets (e.g., IP packets) and are thus only able to detect attacks that are manifested at the network layer (e.g. port scans).

- **Internet Security (IPSec)** —IPSec is the security framework standardized by the IETF as the primary network layer protection mechanism. IPSec consists of two parts; an Authentication Header (AH), whose purpose is to bind the data content of IP frames to the identity of the originator and an Encapsulating Security Payload (ESP) for privacy. The authentication header is intended for use when integrity of information is required but privacy is not. ESP is intended for use where data confidentiality is required. ESP defines two methods (or modes) of encapsulating information. Tunnel mode, when used at an enclave boundary, aggregates traffic flow from site to site and thereby hides end system identification. Transport mode leaves end system identification in the clear and is most advantageous when implemented at the end system.
- **Internet Key Exchange (IKE) Protocol**—IKE was developed by the IETF as a standard for security attribute negotiation in an IP network. It provides a framework for creating security associations between endpoints on an IP network, as well as the methodology to complete the key exchange. IKE is based upon the Internet Security Association Key Management Protocol (ISAKMP) with OAKLEY extensions. The structure of ISAKMP is sufficiently flexible and extensible to allow inclusion of future security mechanisms and their associated algorithms and can be tailored to other networking technologies.
- **Media Encryptors**—Media encryptors protect the confidentiality and integrity of the contents of data storage media. They can also perform a role in maintaining the integrity of the workstation by verifying the Basic Input/Out System (BIOS) and ensuring that configuration and program files are not modified. Media encryptors need to leave some system files unencrypted so that the computer can boot from the hard drive. Most of these files can have their integrity protected by a cryptographic checksum; this will not prevent a tamper attack, but it will alert the user that the data has been altered. However, some system files contain data that changes when the computer is booted; these files cannot be protected. With the exception of some system files, media encryptors encrypt the entire contents of the drive.
- **Packet Filter**—Packet filtering firewalls (also referred to as screening routers) commonly operate at the network layer (Open Systems Interconnection (OSI) Layer 3). These firewalls examine the IP and protocol headers against a set of predefined rules. They can typically filter packets based on host and destination IP address, port number, and the interface. This type of firewall is generally inexpensive, fast, and transparent to the user. However, screening routers generally do not have a very robust auditing capability, nor do they allow the use of strong authentication on incoming connections. The combination of a packet filtering system and another product (authentication server) may provide strong authentication capability.
- **PKI Certificate Management Protocol**—For managing public key material, the Internet community has developed the Internet X.509 Public Key Infrastructure Certificate Management Protocols (CMP). Management protocols are required to support on-line interactions between PKI components. For example, a management protocol might be used between a CA and a client system with which a key pair is associated, or between two CAs that cross-certify each other. At a high level, the set of operations for which management messages are defined can be grouped as follows.

- CA Establishment. When establishing a new CA, certain steps are required (e.g., production of initial CRL, export of CA public key).
  - End Entity Initialization. This includes importing a root CA public key and requesting information about the options supported by a PKI management entity.
  - Certification. Various operations result in the creation of new certificates:
    - Initial registration/certification,
    - Key pair update,
    - Certificate update,
    - CA key pair update,
    - Cross-certification, and
    - Cross-certificate update.
  - Certificate/CRL Discovery Operations. Some PKI management operations result in the publication of certificates or CRLs:
    - Certificate publication, and
    - CRL publication.
  - Recovery Operations. Some PKI management operations are used when an end entity has “lost” its key material.
  - Revocation Operations. Some PKI operations result in the creation of new CRL entries and/or new CRLs.
- SSL—SSL exists just above the transport layer and provides security independent of application protocol although its initial implementation was meant to secure the Hypertext Transfer Protocol (HTTP). This effort has migrated to the IETF as the Transport Layer Security (TLS) protocol. The TLS protocol provides data encryption, server authentication, message integrity, and optional client authentication for a TCP/IP connection. It negotiates the invocation of cryptographic algorithms (from a fixed set) and protects all application layer data.
  - S/MIME—S/MIME is a specification for adding security for e-mail in Multipurpose Internet Mail Extensions format, supporting binary attachments as well as text. It offers authentication and confidentiality. S/MIME uses a hybrid approach to providing security, referred to as a digital envelope. The bulk message is encrypted with a symmetric cipher, a public key algorithm is used for key exchanges and for digital signatures, and X.509 certificates support authentication. S/MIME supports anonymity to the extent that it applies the digital signature first, and then encloses the signature and the original message in an encrypted digital envelope, so that no signature information is exposed to a potential adversary.

The S/MIME specification is currently an Internet draft that recommends three symmetric encryption algorithms: Data Encryption Standard (DES), Triple-DES, and RC2 (a symmetric block cipher with a 40-bit key to meet the U.S. Government export requirements). It also builds on the Public Key Cryptography Standards (PKCS), specifically PKCS #7, providing a flexible and extensible message format for representing the results of cryptographic operations, and PKCS #10, a message syntax for certification requests. The S/MIME specification has been submitted to the IETF in an effort to make it an industry-accepted standard.

- **SOCKS**—This protocol supports application-layer firewall traversal. The SOCKS protocol supports both reliable TCP and User Datagram Protocol (UDP) transport services by creating a shim-layer between the application and transport layers. The SOCKS protocol includes a negotiation step whereby the server can dictate which authentication mechanism it supports. Compliant implementations must support Generic Security Services (GSS)-API and username/password authentication modes.
- **Stateful Packet Filter**—Stateful packet filters look at the same headers as do packet filters, but also examine the content of the packet. In addition, this technology is capable of dynamically maintaining information about past packets or state information. Security decisions can then be based on this state information. Because they have the ability to retain state information, stateful packet filters permit UDP-based services (not commonly supported by firewalls) to pass through the firewall. Thus they are advertised as offering greater flexibility and scalability. Stateful packet filtering technology also allows for logging and auditing and can provide strong authentication for certain services.
- **Trusted Computing Base (TCB)** —A Trusted Computer System is a system that employs sufficient hardware and software assurance measures to allow its use for simultaneous processing of a range of sensitive or classified information. Such a system is often achieved by employing a TCB. A TCB is the totality of protection mechanisms within a computer system, including hardware, firmware, and software, the combination of which is responsible for enforcing a security policy. A TCB consists of one or more components that together enforce a unified security policy over a product or system. The ability of a TCB to correctly enforce a unified security policy depends solely on the mechanisms within the TCB and on system administration personnel's correct input of parameters (e.g., a user's clearance level) related to the security policy.
- **Virus Detectors**—Virus detectors can be employed to protect a network or an individual client. A virus can be considered a special form of intrusion involving the classical Trojan horse attack with the ability to reproduce and spread. The virus is normally considered to be limited to the authorizations of the user who is executing the code, but viruses may also exploit flaws in the network that allow the virus to cause a serious privilege state harm.

## 4.5 Robustness Strategy

### Purpose

The robustness strategy describes a process that, when completed in a later release of the IATF, will provide guidance in assessing the *degree of robustness*. Robustness is defined as the level of security mechanism strength and assurances recommended (considered “good enough”) in an Information Security (INFOSEC) solution. At the current stage of development, the strategy deals primarily with the levels within individual security services and mechanisms, based on information on a given value, in a particular (static) threat environment. As discussed below, this is not a complete answer. The process is not intended to provide an endorsement or credential for specific products, nor is it intended to serve as a “cookbook” answer for the robust-

ness of solutions; rather, it offers security engineering guidance to developers, integrators, and risk managers as input to risk management. Users of the IATF can employ the robustness strategy for:

- Providing guidance to help developers and integrators assess what strength of mechanisms what levels of assurance (in development methodology, evaluation, and testing); and (3) what criteria are recommended for a particular configuration meant to protect information of a particular value, with a specific intelligence life, in a specific, static threat environment;
- Defining product requirements for different customer scenarios (value of information, threat, configuration, etc.) for example as described in the IATF;
- Providing feedback to security requirements developers, decision-makers, customer representatives, customers, etc.;
- Constituting developmental requirements when a security solution does not exist;
- Working with academia to foster research in the network security arena, and to educate future engineers, architects, and users in network security technology; and
- Performing subsequent risk assessments made necessary by reconfiguration of the system/network under review or by a change in threat or value of information.

As technology in general and INFOSEC threats in particular evolve, countermeasures will need to evolve, and with them the corresponding application guidance. This paper is a strategy for the development of a general security mechanism/countermeasure valuation scheme. Rather than directly defining the security requirements, that need to be met, it characterizes the relative strength of mechanisms, that provide security services and provides guidance in selecting these mechanisms.

Trained ISSEs [11] support customer organizations in defining and applying security solutions to address their Information Assurance (IA) needs. Working with a customer from initial contact through solution acceptance, an ISSE helps ensure that the customer's security needs are appropriately identified and that acceptable solutions are developed. Within the context of the IATF robustness strategy, an ISSE helps the customer assess the value of his or her information and assets and the security threat within the operational environment, identifies the security services necessary to provide appropriate protection, and provides guidance on the characteristics of the specific security mechanisms that provide those services.

Multiple applications of the same system/environment but with differently trained ISSEs may result in different guidance, and though all such outcomes would be consistent with the recommended use of the strategy. There is no concept of official compliance with the robustness strategy in terms of approving a solution. It is a strategy, that is an aid to "getting you there" as opposed to a prescriptive solution (where nominal compliance assures acceptability).

## Robustness Strategy Section Overview

The Overview of the General Process section describes the general process, including assumptions and output. Section 4.5.2, *Determining the Degree of Robustness* section presents an approach for determining recommended robustness (strength of mechanism and assurance) levels based on the value of information to be protected and the threat environment. Section 4.5.3,

*Strength of Mechanism*, breaks down security services into supporting mechanisms and identifies corresponding strength levels. The *Level of Assurance* section (Section 4.5.4) discusses related aspects of obtaining assurance. A *Worked Example* (Section 4.5.5) demonstrates how the process would be applied in developing specific guidance. This is followed by a discussion of *Robustness Strategy Evolution* (Section 4.5.6) which provides recommendations for those who would carry on the work outlined in this paper. Then lastly, Section 4.5.7, *Real World Applications*, which demonstrates real world application of the Robustness Strategy.

## 4.5.1 Overview of the General Process

The robustness strategy is intended for application in the context of the development of a security solution and is meant to be consistent with IATF Chapter 3 (Information System Security Engineering), which describes the overall process. An integral part of that process is determining the recommended strength and degree of assurance of proposed security services and mechanisms that become part of the solution set. The strength and assurance features serve as a basis for the selection of the mechanisms and as a means of evaluating the products that implement those mechanisms. This section provides guidance on determining the recommended strength and assurance.

The process should be applied to all components of a solution, both products and systems, to determine the robustness of configured systems as well as of their component parts. It applies to Commercial Off-The Shelf (COTS), Government Off-The-Shelf (GOTS), and hybrid solutions. As indicated above, the process is to be used by security requirements developers, decision-makers, ISSEs, customers, and others involved in the solution life cycle. Clearly, if a solution component is modified, or threat levels or the value of information changes, there must be a reassessment of risk with respect to the new configuration.

Various risk factors, such as degree of damage suffered if the security policy is violated, threat environment, and so on, will be used to guide determination of an appropriate strength, and associated level of assurance for each mechanism. Specifically, the value of information to be protected and the perceived threat environment are used to obtain guidance on the recommended Strength of Mechanism Level (SML) and Evaluation Assurance Level (EAL).

## 4.5.2 Determining the Degree of Robustness

We define the degree of robustness as the level of strength and assurance recommended for potential security mechanism(s). To determine this level for a given security service in a particular application, the customer and the ISSE should consider the value of the information to be protected (in relation to the operational mission) as well as the perceived threat environment. Guidelines for determining these values are provided below. Once a determination has been made regarding the information value and threat environment, the ISSE uses the Robustness Table 4-7 to determine required assurance and strength of mechanism levels.

Note that the robustness strategy focuses specifically on individual security services and mechanisms. When the robustness of an overall network solution is considered, the individual solutions at each layer within the network must also be considered. IA mechanisms can be applied at the host, sub-net, boundary, and backbone levels. Robustness should take into account

the implications of composing layered protection mechanisms and also incorporates an overall assessment of vulnerabilities and residual risks for each layer.

Many customers, in support of their mission, have a need to protect information (or an information system) whose compromise could adversely affect the security, safety, financial posture, or infrastructure of the organization. Five levels of information value have been defined:

- *V1*: Violation of the information protection policy would have negligible adverse effects or consequences.
- *V2*: Violation of the information protection policy would adversely affect and/or cause minimal damage to the security, safety, financial posture, or infrastructure of the organization.
- *V3*: Violation of the information protection policy would cause some damage to the security, safety, financial posture, or infrastructure of the organization.
- *V4*: Violation of the information protection policy would cause serious damage to the security, safety, financial posture, or infrastructure of the organization.
- *V5*: Violation of the information protection policy would cause exceptionally grave damage to the security, safety, financial posture, or infrastructure of the organization.

Similarly, the customer must work with an ISSE to define the threat environment in which the mission will be accomplished. Things to consider when determining the threat to a particular solution include level of access, risk tolerance, expertise, and available resources obtainable by the adversary. These threats should be considered in the context of the system security policy.

The following threat levels were derived from various relevant works (e.g., Security Management Infrastructure (SMI) Task 1 Team, Threat and Vulnerability Model for Information Security, 1997 [12]), and discussions with subject matter experts throughout the Information Systems Security Organization (ISSO). Seven levels of threat have been defined.

- *T1*: Inadvertent or accidental events (e.g., tripping over the power cord).
- *T2*: Passive, casual adversary with minimal resources who is willing to take little risk (e.g., listening).
- *T3*: Adversary with minimal resources who is willing to take significant risk (e.g., unsophisticated hackers).
- *T4*: Sophisticated adversary with moderate resources who is willing to take little risk, e.g., organized crime, sophisticated hackers, international corporations.
- *T5*: Sophisticated adversary with moderate resources who is willing to take significant risk (e.g., international terrorists).
- *T6*: Extremely sophisticated adversary with abundant resources who is willing to take little risk (e.g., well-funded national laboratory, nation-state, international corporation).
- *T7*: Extremely sophisticated adversary with abundant resources who is willing to take extreme risk, (e.g., nation-states in time of crisis).

After a determination is made regarding the value of the information to be protected and the threat environment, the ISSE can provide guidance on how strong the security mechanism should be and what assurance activities that should be performed. Table 4-7 indicates the minimal recommended SML and EAL[6] or providing protection of information or information

systems of a given value (V1toV5) against a given threat level (T1toT7). Section 4.5.3 (Strength of Mechanism) defines the SMLs and Section 4.5.4 (Level of Assurance) defines the EALs.

**Table 4-7 Degree of Robustness**

Information Value	Threat Levels						
	T1	T2	T3	T4	T5	T6	T7
<b>V1</b>	SML1 EAL1	SML1 EAL1	SML1 EAL1	SML1 EAL2	SML1 EAL2	SML1 EAL2	SML1 EAL2
<b>V2</b>	SML1 EAL1	SML1 EAL1	SML1 EAL1	SML2 EAL2	SML2 EAL2	SML2 EAL3	SML2 EAL3
<b>V3</b>	SML1 EAL1	SML1 EAL2	SML1 EAL2	SML2 EAL3	SML2 EAL3	SML2 EAL4	SML2 EAL4
<b>V4</b>	SML2 EAL1	SML2 EAL2	SML2 EAL3	SML3 EAL4	SML3 EAL5	SML3 EAL5	SML3 EAL6
<b>V5</b>	SML2 EAL2	SML2 EAL3	SML3 EAL4	SML3 EAL5	SML3 EAL6	SML3 EAL6	SML3 EAL7

Using an applicable Capability Maturity Model (CMM), Capability Level 2 or equivalent is recommended for EALs 1to3 and a Capability Level 3 or equivalent for EALs 4-7.

The Systems Security Engineering Capability Maturity Model (SSE-CMM) is one example of an applicable CMM. A capability maturity model describes the stages through which processes advance as they are defined, implemented, and improved.<sup>1</sup> The SSE-CMM is designed to support a host of improvement activities, including self administered appraisals or internal appraisals augmented by experts (e.g., ISSEs) from inside or outside of the organization.<sup>2</sup>

The ISSE, working with the customer, would apply the SSE-CMM (or another applicable CMM) as a baseline capability. The assessment of compliance is still left to the discretion of the customer. Reasonable justification is still necessary and it should be denoted that acquisition personnel need to be knowledgeable about the CMM used.

### 4.5.3 Strength of Mechanism

SML is presented by a series of tables focusing on specific security services. The strategy is still being formulated, and the tables are not considered complete or adequately refined. There are a number of additional security mechanisms that are not detailed in the tables but that may be appropriate for providing some security services. Further, the strategy is not intended, by itself to provide adequate information for the selection of the desired (or sufficient) mechanisms for a particular situation. As indicated earlier, an effective security solution will only result from

<sup>1</sup> System Security Engineering Capability Maturity Model Description document

<sup>2</sup> System Security Engineering Capability Maturity Model Summary



the proper application of ISSE skills to specific operational and threat situations. The strategy does offer a methodology for structuring a more detailed analysis. The security services itemized in these tables have several related supporting security services that may result in recommendations for inclusion of additional security mechanisms and techniques.

For each service, recommended guidance on each SML levels is given for a variety of mechanisms that provide the overall service. In some cases, a group of mechanisms will be required to provide the necessary protection. It should also be noted that an ISSE, in conjunction with a customer, could decide to use a stronger or weaker mechanism than is recommended, depending on the environment. It is the intent of the strategy to ensure that mechanisms across services at the same strength level provide comparable protection, in that they counter equivalent threats. The selection of mechanism(s) from service tables is an independent event, in the sense that one mechanism does not necessarily require others. Higher strength mechanisms do not necessarily contain features of lower strength mechanisms (i.e., security functions do not necessarily accumulate at higher strength levels). Table entries are preliminary estimates based on consultation with subject matter experts and are likely to be revised based on technology evolution, threat assessment, and costing development.

The strength referred to below is a *relative* measure of the effort (cost) required to defeat the mechanism and is not necessarily related to the cost of implementing such countermeasures. All things being equal, (especially cost), the highest strength mechanism should always be chosen. Three SMLs are defined:

- *SML1* is defined as basic strength or good commercial practice. It is resistant to the unsophisticated threat (roughly comparable to the T1toT3 threat levels) and is used to protect low-value data. Examples of countered threats might be door rattlers, ankle biters, inadvertent errors.
- *SML2* is defined as medium strength. It is resistant to the sophisticated threat (roughly comparable to the T4toT5 threat levels) and is used to protect medium-value data. It would typically counter a threat from an organized effort (e.g. an organized group of hackers).
- *SML3* is defined as high strength or high grade. It is resistant to the national laboratory or nation-state threat (roughly comparable to the T6toT7 threat levels) and is used to protect high-value data. An example is an extremely sophisticated, well-funded technical laboratory or a nation-state adversary.

Based on these definitions, the customer and ISSE will apply their knowledge of the specific operational and threat situation to determine what strength of mechanism is recommended for each of the mechanisms listed in the following sections.

### 4.5.3.1 Mechanisms Supporting Security Management

Recommended mechanisms for establishing needed security management are depicted in Table 4-8. The degree of awareness and control with respect to the following will identify the SML target.

- *Compromise recovery*, in addition to achieving a secure initial state, secure systems must have a well-defined status after failure, either to a secure failure state or via a recovery procedure to a known secure state.
- *Poor system administration* is a leading cause of security weaknesses and vulnerabilities. It is the first line of defense in enforcing the security policy. (See IATF Chapter 3 Information Systems Security Engineering) for more information on system security administration.
- *Training* is what operators and users need to obtain to learn about security features and system operation. Knowledgeable users are more likely to exercise due care in protecting information assets (increased risk of insider attack is dealt with via by means of personnel security).
- *The Operational Security (OPSEC)* process is a coordinated, multidisciplinary five-step activity involving identification of critical information, threat identification and analysis, vulnerability identification and analysis, risk assessment, and adoption of countermeasures. Each use of the process is tailored to a specific activity of concern, which is examined for potential disclosure to specific adversaries, upon which to base directly pertinent countermeasures. Consult with the Interagency Operation Support Staff for consideration of individual cases.
- *Trusted distribution* is a calculated/controlled method of distributing security-critical hardware, software, and firmware components. It protects of the system from modification during distribution and detects any changes.
- *Secure operations* is the level of standard operating procedures needed to provide security given the classification, sensitivity, and criticality of the data and resources being handled or managed. This includes security doctrine.
- *Mechanism management*, certain security mechanisms (e.g., cryptographic algorithms) have ancillary support needs (e.g., key management).

**Table 4-8 Security Management Mechanisms**

	<b>Compromise Recovery</b>	<b>System Administration</b>	<b>Training</b>	<b>OPSEC</b>	<b>Trusted Distribution</b>	<b>Secure Operations</b>	<b>Mechanism Management</b>
<b>SML1</b>	Informal plan	See Ch. 4 counter-measures	Training available at user discretion	Implement OPSEC at user's discretion	Direct vendor purchase	Informal plan of operation	Procedural, user's discretion
<b>SML2</b>	Detailed plan that is reviewed and approved	See Ch. 4 counter-measures	Formal training plan	OPSEC training required; implement at user's discretion	Certificate of authenticity, virus scan, validation	Formal plan of operation	Procedural, reminders, user's discretion
<b>SML3</b>	Detailed plan that is reviewed and approved	See Ch. 4 counter-measures	Knowledge/skill certification required	OPSEC training required, implementation required	Protective packaging, checksums, validation suite	Detailed, formal plan of operation	Automated support

### 4.5.3.2 Mechanisms Supporting Confidentiality

Confidentiality is the protection of information against disclosure to unauthorized entities or processes. Possible security mechanisms for this security service are depicted in Table 4-9. These mechanisms can be obtained individually or in combination.

- *If cryptographic algorithm* is chosen, some of the factors that must be considered are the management of keying material and the effective length of the key, which includes the strength of the underlying cryptographic algorithm. Effective key length is defined as the nominal key length, reduced by the effect of any known attacks against the cryptographic algorithm (assuming correct implementation). The supporting KMI [9] categories are defined in Chapter 8 (Supporting Infrastructures).
- *Physical security* includes tangible security mechanisms such as guards, locks, and fences. The idea is to build a physically secure enclave, providing guards and high walls.
- *Technical security* is a protection mechanism for hardware. Tampering is the unauthorized modification that alters the proper functioning of an information security device or system in a manner that degrades the security or functionality it provides. Anti-Tamper mechanisms detect such alterations. TEMPEST is the investigation, study, and control of compromising emanations from telecommunications and Automated Information System (AIS) equipment.
- *Anonymity* is the desire for a user to remain unknown during a virtual transaction. Some applications requiring anonymity might be Internet voting and Internet cash. This area is relatively immature and is currently addressed by the Transmission Security (TRANSEC)[10] and cover & deception disciplines. TRANSEC mechanisms provide various degrees of covertness to prevent detection, identification and exploitation. Cover and deception can be provided through such mechanisms as anonymous remailers, “onion routing”, or “web anonymizers.” Cover and deception currently has no differentiated levels.

**Table 4-9 Confidentiality Mechanisms**

	Cryptographic Algorithm		Physical Security	Technical Security		Anonymity	
	Effective Key Length	Key Management		Anti tamper	TEMPEST	TRANSEC	Cover & Deception
<b>SML1</b>	40+ bits symmetric key length, 80+ exponent 512+ modulus public key length	SMI Cat X, 80+ exponent 512+ modulus public key length, 80+ hash key length	Comparable to [7]	[6] level 1 or 2	comply with applicable EMI/EMC FCC standards or portions of [8]	Low power unit	TBD
<b>SML2</b>	80+ bits symmetric key length, 160+ exponent 1024+ modulus public key length	SMI Cat Y, 160+ exponent 1024+ modulus public key length, 160+ hash key length	Comparable to [7]	[6] level 3 or 4	[8]	Commercial spread spectrum signal techniques	TBD
<b>SML3</b>	Because of the complicated nature of this level, please consult with a qualified ISSE. <sup>3</sup>	SMI Cat Z, also consult with a qualified ISSE. <sup>3</sup>	Comparable to [7]	[6] level 4 or better	[8]	cryptographic spread spectrum signal techniques	TBD

### 4.5.3.3 Mechanisms Supporting Integrity

In Table 4-10 there are four mechanisms that will help in ensuring integrity, either singly or in combination with others. When taken in the context used here, integrity, as a security service, means the protection of information against undetected, unauthorized modification, or undetected destruction of information.

- A cryptographic algorithm in an error extension mode will emphasize the error and should be used in conjunction with a detection mechanism (e.g., parity or human review).
- Physical security is described in Table 4-9.
- Signature/Checksum provides data integrity by digitally signing data. Typically, the data requiring protection is used to calculate a smaller value, such as a parity, checksum, or hash. This value can then be digitally signed.

<sup>3</sup> DoD users should consult with a National Security Agency ISSE. Other government users are directed to contact an ISSE at the National Institute of Standards and Technology for guidance in this area. Non-government users should consult with a qualified ISSE, or equivalent representative within their organization.

- Redundancy is the availability of multiple methods to obtain the same information.

**Table 4-10 Integrity Mechanisms**

	Cryptographic Algorithm		Physical Security	Signature Checksum	Redundancy
	Effective Key Length	Key Management			
<b>SML1</b>	40+ bits symmetric key length, 80+ exponent 512+ modulus public key length	SMI Cat., 80+ exponent 512+ modulus public key length, 80+ hash key length	comparable to [7]	Parity, or commercial checksum, hash and signature with SML1 algorithm	Not applicable
<b>SML2</b>	80+ bits symmetric key length, 160+ exponent 1024+ modulus public key length	SMI Cat, 160+ exponent 1024+ modulus public key length, 160+ hash key length	comparable to [7]	Cryptographic checksum, hash, and signature with SML2 algorithm	Redundant data path with 100% correct comparison
<b>SML3</b>	Due to the complicated nature of this level, please consult with a qualified ISSE. <sup>4</sup>	SMI Cat, also consult with a qualified ISSE. <sup>4</sup>	comparable to [7]	Cryptographic checksum, hash and signature with SML3 algorithm	Multiple data paths with 100% correct comparison

### 4.5.3.4 Mechanisms Supporting Availability

Availability is also known as service assurance. To ensure availability of data, the system must employ both preventive and recovery mechanisms. This security service is quantified in Table 4-11 and can be obtained through a combination of the services as appropriate for the applications.

- TRANSEC is used to overpower potential jammers. A strong enough signal is provided for this anti-jam capability. TRANSEC can also be used to hide a signal to avoid jamming. (Note that, because of the real-time nature of exploitation, it may not be necessary to use an SML3 algorithm strength to meet the SML3 level for this mechanism).
- Anti-tamper mechanism is described in Table 4-9.
- Physical security is described in Table 4-9.
- Redundancy or redundant paths should be available to allow information flow without violating the site security policy. Such information flow might include bypassing any problem areas, including congested servers, hubs, cryptography and so on.
- Data recovery is the ability to recover data that might otherwise be unavailable due to the loss of key, storage media, etc.

<sup>4</sup> DoD users should consult with a National Security Agency ISSE. Other government users are directed to contact an ISSE at the National Institute of Standards and Technology for guidance in this area. Non-government users should consult with a qualified ISSE or an equivalent representative within their organization.

**Table 4-11 Availability Mechanisms**

	TRANSEC	Anti-Tamper	Physical Security	Redundancy	Data Recovery
<b>SML1</b>	High power	level 1 or 2 [4]	comparable to [7]	Bypass channel available	Informal archival plan, user backs up own key or data
<b>SML2</b>	Commercial spread spectrum signal techniques	level 3 or 4 [4]	comparable to [7]	Backup data path, hot spare	Formal archival plan, central back-ups
<b>SML3</b>	Cryptographic spread spectrum signal techniques	level 4 or better [4]	comparable to [7]	Multiple data paths, multiple hot spares	Formal archival plan, central, offsite back-ups

### 4.5.3.5 Mechanisms Supporting Identification and Authentication (I&A)

I & A is required for effective access control. There usually is a need for a process that enables recognition of an entity within or by an AIS. Along with that, a security measure designed to establish the validity of a transmission, message, or originator or a means of verifying an individual's eligibility to receive specific categories of information is needed. These are the attributes of I&A that are listed in Table 4-12. We categorize these attributes as follows.

- *Identification or System Identification (SID)* in particular is one way in which a system might recognize the “entity” (which may be a person requesting authentication). Biometrics might be used to identify a living person.
- *Human-to-machine authentication* could utilize alphanumeric phrases, like passwords, Personal Identification Numbers (PIN), or challenge/response exchanges that are memorized by a human or used with a token calculator (e.g. challenge/response). Also, physical devices, such as hardware tokens, have this utility (e.g., a credit card-type physical entity).
- *Peer-to-peer authentication* can utilize certificates that identify and authenticate the entities. Along with the certificate is the similar SML cryptographic algorithm that “binds” it to the entity with a digital signature. Authentication is provided by a trusted third party (a separate, but knowledgeable entity). Within this area, one could use a cryptographic algorithm (as discussed under confidentiality above), and personnel security policy, where a security clearance is obtained for a particular person to reduce the risk of an insider attacking the system.

**Table 4-12 Identification and Authentication Mechanisms**

	Identification		Human-to-Machine Authentication		Peer-to-Peer Authentication			
	System IDs	Bio-metrics	Passwords PINS Challenge/Response	Tokens	Certificates	Cryptographic Algorithm		Personnel Security
						Effective Key Length	Key Management	
<b>SML1</b>	Uniqueness	Not applicable	Have one	Badge/key static	bind w/SML1 cryptographic algorithm	40+ bits symmetric key length, 80+ exponent 512+ modulus public key length	SMI Cat. X, 80+ exponent 512+ modulus public key length, 80+ hash key length	Commercial hiring practices
<b>SML2</b>	Uniqueness and minimum character length	Use one Biometric	Minimum effective length - TBD	Memory device, updated periodically	bind w/SML2 cryptographic algorithm	80+ bits symmetric key length, 160+ exponent 1024+ modulus public key length	SMI Cat Y, 160+ exponent 1024+ modulus public key length, 160+ hash key length	Equivalent of Secret clearance
<b>SML3</b>	Uniqueness and minimum character length, minimum distance (e.g., Hamming)	Use one Biometric with a liveness test	Minimum effective length - TBD	CIK, updated every time	bind w/SML3 cryptographic algorithm	Because of the complicated nature of this level, please consult with a qualified ISSE. <sup>5</sup>	SMI Cat Z, also consult with a qualified ISSE. <sup>5</sup>	Equivalent of Top Secret clearance

### 4.5.3.6 Mechanisms Supporting Access Control

Beyond I&A, access control can be thought of as a “super service” encompassing all security services. In the context of network security, access control is concerned with limiting access to networked resources (hardware and software) and data (stored and communicated). The primary goal here is to prevent unauthorized use, and unauthorized disclosure or modification of data by unauthorized entities. A secondary goal is to prevent an availability attack (e.g., denial-of-service attack). Several mechanisms that can be used to help provide the access control service are shown in Table 4-13 and include the following parameters.

- *Anti-tamper* is described under Confidentiality (Table 4-9).

<sup>5</sup> DoD users should consult with a National Security Agency ISSE. Other government users are directed to contact an ISSE at the National Institute of Standards and Technology for guidance in this area. Non-government users should consult with a qualified ISSE, or an equivalent representative within their organization.

- *Mandatory Access Control (MAC)* is where authorized access to data is automatically imposed by the system through the use of labels and binding the labels to the data associated with it. When implementing MAC, there is a concern with the integrity of the label itself and the strength of binding of the label to the data. In other words, if SML2 is required for MAC, the integrity of the label must be provided with SML2, and the function (possibly a cryptographic algorithm) binding the label to the data must also be SML2. Other implementation concerns include making the labeling non-bypassable and fail-safe.
- *Discretionary Access Control (DAC)* is different from MAC in that the owner of the data to be accessed (versus the machine) can choose who can and cannot be authorized access to the data. For SML1, this is comparable to setting UNIX permission bits (owner/group/world) to grant access. For SML2 and 3, using ACLs further refines the mechanism. ACLs can be more specific to allow certain identities access to information, (e.g. specific users within a group can be granted access). Again, DAC mechanisms should be non-bypassable (only “changeable” by the owner of the data), fail-safe, and possess the same SML level of integrity associated with the level of DAC required.
- *Certificates* are described under I&A (Table 4-12).
- *Personnel security* is described under I&A (Table 4-12).

**Table 4-13 Access Control Mechanisms**

	Anti-Tamper	Mandatory Access Control	Discretionary Access Control	Certificates	Personnel Security
<b>SML1</b>	level 1 or 2 [4]	Not applicable	Comparable to Unix permission bits	bind w/SML1 cryptographic algorithm	Commercial hiring practices
<b>SML2</b>	level 3 or 4 [4]	Labels bound to data having integrity and binding function both at the SML2 level	ACLs	bind w/SML2 cryptographic algorithm	Equivalent of Secret clearance
<b>SML3</b>	level 4 or better [4]	Labels bound to data having integrity and binding function both at the SML3 level	ACLs	bind w/SML3 cryptographic algorithm	Equivalent of Top Secret clearance

### 4.5.3.7 Mechanisms Supporting Accountability

Accountability can be considered a special case of non-repudiation. The accountability security service is basically holding any network entity responsible for its actions on that network. Mechanisms, that can be used to provide the security service of accountability are shown in Table 4-14, and discussed below.

- When implementing the *audit* mechanism, the following components should be considered.
  - What is being audited and relevant events that are detected.
  - How the audit (detected) data is protected, analyzed, and reported.



- What the reaction strategy is to the audit data analysis and reporting.

These components should be considered for each SML level, and in SML2 and 3, should be detailed in a plan. As with all mechanisms, consideration should be given to noncircumvention or “non-bypassability” and the effects of failure.

- *Intrusion detection* is still in relative infancy. Intrusion detection is that mechanism that monitors a network and detects either (1) known attacks being mounted against the system or (2) differences in a profiled use of the system. Several aspects associated with an intrusion detection mechanism for examples, whether it is static [SML1] set up to filter only on known attacks and profiles); dynamic [SML2] set up to filter on known attacks and profiles but updateable perhaps through software downloads), or dynamically adaptable [SML3] this adds the aspect of “artificial intelligence” in which the system learns new profiles based on usage). Depending on the SML level, a reaction mechanism to a detected intrusion must be either informally (SML1) or formally (SML2 and 3) detailed and implemented.
- *I&A* is described under I&A (Table 4-12).

**Table 4-14 Accountability Mechanisms**

	<b>Audit</b>	<b>Intrusion Detection</b>	<b>I&amp;A</b>
<b>SML1</b>	informal reaction mechanism	static system with informal reaction mechanism	see I&A table for SML1
<b>SML2</b>	formal reaction plan and strategy	dynamic system with formal reaction mechanism	see I&A table for SML2
<b>SML3</b>	formal reaction plan and strategy	dynamic, adaptive system with formal reaction mechanism	see I&A table for SML3

### 4.5.3.8 Mechanisms Supporting Non-Repudiation

The security service of non-repudiation provides a method by which the sender of data is provided with proof of delivery and the recipient is assured of the sender’s identity, so that neither can later deny processing the data. It is quantified in Table 4-15 and can be obtained by a combination of these mechanisms as appropriate for the applications:

- *Signature* is used to digitally sign data in such a way that only the sender and receiver could have respectively sent and received the message. The sender signs the original data to prove that it was sent. The receiver signs a receipt as proof of receipt of the original data. Validation of these signatures is always required.
- *Trusted Third Party* is used to prearrange a method by which a third party may receive the information from the sender and transmit/send it to the receiver in a way that ensures that the sender and receiver are confident that they are communicating with the correct party.
- *Accountability* is described under Accountability (Table 4-14)
- *I&A* is described under the I&A (Table 4-12).
- *Archive* is the ability to store data so that it can be recovered if necessary.

**Table 4-15 Non-Repudiation Mechanisms**

	<b>Signature</b>	<b>Trusted Third Party</b>	<b>Accountability</b>	<b>I&amp;A</b>	<b>Archive</b>
<b>SML1</b>	sign with SML1 cryptographic algorithm	see I&A Table for SML1 Personnel Security	see Accountability table for SML1	see I&A table for SML1	informal archival plan, user backs up own key or data
<b>SML2</b>	sign with SML2 cryptographic algorithm	see I&A Table for SML2 Personnel Security	see Accountability table for SML2	see I&A table for SML2	formal archival plan, central back- ups
<b>SML3</b>	sign with SML3 cryptographic algorithm	see I&A Table for SML3 Personnel Security	see Accountability table for SML3	see I&A table for SML3	formal archival plan, central, offsite back-ups

## 4.5.4 Level of Assurance

The discussion addressing the need for an overall system security solution view of strength of mechanism is also relevant for the level of assurance. Again, while an underlying methodology is offered, a real solution can only be deemed effective after a detailed analysis activity that considers the specific operational and threat situations and the system context for the solution.

Assurance is the measure of confidence in claims made and that the security features and architecture of an automated information system appropriately mediate access and enforce the security policy. The assurance measures listed here are from the Common Criteria [6].

The Common Criteria provide assurance through active investigation. Active investigation is an evaluation of the actual product or system to determine its actual security properties. The Common Criteria philosophy assumes that greater assurance results come from greater evaluation effort in scope, depth, and rigor. This leads to the seven EALs (EAL1 through EAL 7) described below:

- *EAL 1, Functionally Tested*, is applicable where some confidence in correct operation is required, but the threats to security are not viewed as serious. It is of value where independent assurance is required to support the contention that due care has been exercised with respect to the protection. An example is the protection of personal information.
- *EAL 2, Structurally Tested*, requires the cooperation of the developer in terms of the delivery of design information and test results, but should not demand more effort (or substantially increased cost or time) than is consistent with good commercial practice. It is applicable where a low to moderate level of independently assured security is required in the absence of an available development record. An example is securing legacy systems, or where access to the developer is limited.
- *EAL 3, Methodically Tested and Checked*, permits a conscientious developer to gain maximum assurance from positive security engineering at the design stage without substantial alteration of existing sound development practices. It is applicable where a moderate level of independently assured security is required.

- *EAL 4, Methodically Designed, Tested and Reviewed*, permits a developer to gain maximum assurance from positive security engineering based on good commercial development practices which, though rigorous, do not require substantial specialist knowledge, skills, and other resources. This is the highest level at which it is likely to be economically feasible to retrofit to an existing product line. It is applicable in those circumstances where a moderate to high level of independently assured security in conventional products is required, and where developers or users are prepared to incur additional security-specific engineering costs.
- *EAL 5, Semi-Formally Designed and Tested*, permits a developer to gain maximum assurance from security engineering based on rigorous commercial development practices supported by moderate application of specialized security engineering techniques. It is applicable where a high level of independently assured security in a planned development is required along with a rigorous development approach.
- *EAL 6, Semi-Formally Verified Design and Tested*, permits developers to gain high assurance from application of security engineering techniques to a rigorous development environment to protect high value assets against significant risks. It is applicable to the development of security products to be used in high-risk situations.
- *EAL 7, Formally Verified Design and Tested*, is applicable to the development of products to be used in extremely high risk situations and/or where the high value of the assets justifies the higher costs. Realistically, it is limited to products with tightly focused functionality that is amenable to extensive formal analysis.

The assurance levels are a composition of the following assurance classes: Configuration Management, Delivery and Operation, Development, Guidance Documents, Life-Cycle Support, Tests, and Vulnerability Assessments. These classes incorporate the concepts of correct implementation, “non-bypassable” mechanisms, failure to a secure state, secure startup, and others.

In addition to those addressed in the Common Criteria, there are other assurance tasks that the Common Criteria do not discuss, including Failure Analysis and Test, TEMPEST Analysis and Test, and Tamper Analysis and Test. If these apply to a particular product or system, then they should be added to the requirements of the appropriate EALs.

## 4.5.5 A Worked Example

### Assumptions

- Security evaluation is a necessary part of solution development.
- A trained ISSE (or equivalent) is the strategy consumer.

The methodology for correct employment of the robustness strategy is as follows.

- The responsible customer party knows, and has appropriately documented, mission objectives, concept of operation, value of information to be protected, threat/environment context, and security policy.

- A solution is then engineered (per IATF Chapters 5 through 9 [Defend the Network and Infrastructure, Defend the Enclave Boundary/External Connections, Defend the Computing Environment, Supporting Infrastructures, and Information Assurance for the Tactical Environment], providing guidance on the security mechanisms required.
- Risk factors (e.g., degree of damage if security policy is violated, threat environment) will be used to help determine the appropriate strength and associated level of assurance for each mechanism from the set of security service tables. The risk addressed is the residual risk, not the overall (or initial) risk, what remains after other countermeasures have been applied, and what would be the target of doctrine if additional security measures were not taken. For example, a system-high workstation in a “secure” office setting has a different residual risk from that same workstation operating in a public environment.
- Working with an ISSE, the customer will then select COTS/GOTS product(s) providing the necessary strength and assurance.
- Evaluate the system and highlight the residual risk.

### 4.5.5.1 Example One

The following illustration uses an abbreviated example of the media protection portion of the IATF Remote Access (Section 6.2), Secret Dial-in Case, to demonstrate how the robustness strategy would typically be used in conjunction with other guidance sections of the IATF. No attempt was made to consider an actual customer’s needs or an actual recommended solution.

In this example, the customer will be processing Secret data at a Continental U.S. (CONUS) site (perhaps in a work-at-home or temporary duty (TDY) situation) on a remote access dial-in system. The customer is required to protect this data and feels the threat to the data is primarily from *adversaries*:

- *Who have minimal resources at their disposal* (i.e., they have enough money or contacts so that they can get someone to steal the laptop out of a house/hotel room)
- *Who are willing to take significant risk* (i.e., if the person gets caught stealing, the adversaries are willing to be prosecuted; or know that if the thief gets caught the theft will not be traced back to them).

For this example, a media encryptor is recommended to provide confidentiality of the customer’s Secret data on the hard drive of the remote computer. Because the data is Secret, according to the current classification manual, compromise of that data would *cause serious damage to the security of the United States*. Based on the situation described here, the customer, in conjunction with the ISSE, determines that the value of his information is at the *V4* level (violation of the information protection policy would cause serious damage to the security, safety, financial posture, and/or infrastructure of the organization), and the perceived threat is at the *T3* level (adversary with minimal resources who is willing to take significant risk). Using the Degree of Robustness Table presented in Table 4-7 (as depicted in Table 4-16), the minimum SML and EAL recommended is *SML2* and *EAL3* based on the threat and information levels.

**Table 4-16 Example Depicting Use of Degree of Robustness Table**

Information Value	Threat Levels						
	T1	T2	T3	T4	T5	T6	T7
<b>V1</b>	SML1 EAL1	SML1 EAL1	SML1 EAL1	SML1 EAL2	SML1 EAL2	SML1 EAL2	SML1 EAL2
<b>V2</b>	SML1 EAL1	SML1 EAL1	SML1 EAL1	SML2 EAL2	SML2 EAL2	SML2 EAL3	SML2 EAL3
<b>V3</b>	SML1 EAL1	SML1 EAL2	SML1 EAL2	SML2 EAL3	SML2 EAL3	SML2 EAL4	SML2 EAL4
<b>V4</b>	SML2 EAL1	SML2 EAL2	<b>SML2</b> <b>EAL3</b>	SML3 EAL4	SML3 EAL5	SML3 EAL6	SML3 EAL6
<b>V5</b>	SML2 EAL1	SML2 EAL3	SML3 EAL4	SML3 EAL5	SML3 EAL6	SML3 EAL6	SML3 EAL7

For our example, and as a result of applying the IATF guidance, the ISSE together with the customer determined that confidentiality and security management services are recommended. The user of the remote access dial-in system will want to keep the Secret data on the laptop inaccessible while in storage. Not only does the data need to be encrypted on the media, but also the system needs to be operated in a secure manner; furthermore, the issue of how to recover the data if it is compromised needs to be addressed. The ISSE and customer together decide that media encryption will be one mechanism used. From the discussions above, a media encryptor of strength SML2 should be considered.

Once the security service has been selected, confidentiality in this case, then the mechanism should be chosen from the columns of that table. In this case, the mechanism chosen is cryptographic algorithm. This mechanism has been chosen because it was the cheapest, simplest, most practical to implement. Physical security was not chosen because it was impossible to uniformly apply, in a timely manner, at different remote sites without knowing all the sites in advance. Technical security was not chosen because of the wide variety of COTS laptops, which are currently not built with technical security countermeasures. Using the Confidentiality Mechanisms Table, as depicted in Table 4-17, the implementation should look for a cryptographic algorithm capability with an effective key length of 80+ bits, supported by a KMI/PKI providing the strength under category “Y” as further described in Chapter 8-1 (KMI/PKI).

**Table 4-17 Example Depicting Use of Confidentiality Mechanisms Table**

	Cryptographic Algorithm		Physical Security	Technical Security		Anonymity	
	Effective Key Length	Key Management		Anti-Tamper	TEMPEST	TRANSEC	Cover
<b>SML1</b>	40+ bits symmetric key length, 80+ exponent 512+modulus public key length	SMI Cat X, 80+exponent 512+ modulus public key length, 80+ hash key length	Comparable to [7]	level 1 or 2 [4]	comply with applicable EMI/EMC FCC standards or portions of [8]	low power unit	TBD
<b>SML2</b>	80+ bits symmetric key length, 160+ exponent 1024+ modulus public key length	SMI Cat, 160+ exponent 1024+ modulus public key length, 160+ hash key length	Comparable to [7]	level 3 or 4 [4]	[8]	commercial spread spectrum signal techniques	TBD
<b>SML3</b>	Because of to the complicated nature of this level, please consult with a qualified ISSE. <sup>6</sup>	SMI Cat Z, also consult with a qualified NSA ISSE. <sup>6</sup>	Comparable to [7]	level 4 or better [4]	[8]	cryptographic spread spectrum signal techniques	TBD

Because the remote access dial-in users will not have direct access to their system administrator or support services, the customer and ISSE found that the need for the security management mechanisms of training and secure operations was of paramount importance and should be supplied at the SML3 level. Similarly, because of the “remote” use of the system, they thought that compromise might be more likely; and, therefore, the compromise recovery mechanism was also of paramount importance and should be addressed at the SML3 level. Further, because of the value of the information and the threat to the information, the components should be characterized as methodically tested and checked, consistent with the Common Criteria EAL3. (Note that this depicts a situation where the initial SML and EAL recommendations from the strategy were considered inadequate, and were thus increased, presumably based on a detailed analysis of the situation.) Table 4-18 depicts how the Security Management Mechanisms table would typically be used.

Note that when using the tables in this section, not all columns must be used, and various SML levels may be employed as needed for the specific mechanism under question. In the media encryption example, it may be determined that security management mechanisms are of paramount importance; therefore, *SML3* will be chosen while confidentiality mechanisms to protect data may be adequate with a *SML2* cryptographic algorithm.

<sup>6</sup> DoD users should consult with a National Security Agency ISSE. Other government users are directed to contact an ISSE at the National Institute of Standards and Technology for guidance in this area. Non-government users should consult with a qualified ISSE, or equivalent representative within their organization.

**Table 4-18 Example Depicting Use of Security Management Mechanisms Table**

	<b>Compro- mise Recovery</b>	<b>System Admin- istration</b>	<b>Training</b>	<b>OPSEC</b>	<b>Trusted Distribution</b>	<b>Secure Operations</b>	<b>Mechanism Manage-ment</b>
<b>SML1</b>	informal plan	See Ch. 4 countermeasures	training available at user discretion	Implement OPSEC at user's discretion	direct vendor purchase	informal plan of operation	procedural, user's discretion
<b>SML2</b>	detailed plan that is reviewed and approved	See Ch. 4 countermeasures	formal training plan	OPSEC training required, implement at user's discretion	certificate of authenticity, virus scan, validation	formal plan of operation	procedural, reminders, user's discretion
<b>SML3</b>	detailed plan that is reviewed and approved	See Ch. 4 countermeasures	Knowledge/skill certification required	OPSEC training required, implement OPSEC required	protective packaging, checksums, validation suite	detailed, formal plan of operation	automated support

### 4.5.5.2 Example Two

A second example of the use of the strategy is where a Sensitive Compartmented Information Facility (SCIF) is employed as physical protection. Very different security mechanisms would likely be chosen to protect the information. If a DoD system is processing Top Secret data (V5), and the threat is very high (T6), one would normally apply rigorous SML and EAL levels for this solution. However, because the SCIF is used (and there is no connectivity outside the SCIF), the confidentiality requirement is mostly satisfied by physical security at SML3 level. The access control requirement may also be satisfied by personnel security at SML3 level. The residual risk in the areas of confidentiality and access control may be mitigated by additional mechanisms at the SML1 level. This example shows the importance of layering security mechanisms to reduce risk.

### 4.5.5.3 Example Three

A third example involves a corporation with a large Intranet that processes only unclassified data. The corporation has stringent legal requirements to protect its data from unauthorized access or modification. The corporation maintains a large heterogeneous network with Internet access protected by firewalls. All data requiring legal protection is maintained in isolated subnets and is not available to authorized users via the network. Offline stand-alone access is required to view the protected data. The security objective is to upgrade the network to allow the protected data to be securely accessible by all authorized users. Although the data being processed is unclassified but it must be protected from unauthorized access. Using the applicable capability maturity model, a Capability Level 2 or equivalent is recommended. Taking all this into consideration, the customer along with the ISSE determined that the information is at the V3 level (violation of the information protection policy would cause some damage to the security safety, financial posture, and/or infrastructure of the organization), and the perceived

threat is at the *T4* level (sophisticated hackers, international corporations). Using the Degree of Robustness Table presented in Table 4-7 (as depicted in Table 4-19), the minimum SML and EAL recommended is SML2 and EAL3 based on the threat and information levels.

**Table 4-19 Example Depicting Use of Degree of Robustness Table**

Information Value	Threat Levels						
	T1	T2	T3	T4	T5	T6	T7
<b>V1</b>	SML1 EAL1	SML1 EAL1	SML1 EAL1	SML1 EAL2	SML1 EAL2	SML1 EAL2	SML1 EAL2
<b>V2</b>	SML1 EAL1	SML1 EAL1	SML1 EAL1	SML2 EAL2	SML2 EAL2	SML2 EAL3	SML2 EAL3
<b>V3</b>	SML1 EAL1	SML1 EAL2	SML1 EAL2	SML2 EAL3	SML2 EAL3	SML2 EAL4	SML2 EAL4
<b>V4</b>	SML2 EAL1	SML2 EAL2	SML2 EAL3	SML3 EAL4	SML3 EAL5	SML3 EAL5	SML3 EAL6
<b>V5</b>	SML2 EAL2	SML2 EAL3	SML3 EAL4	SML3 EAL5	SML3 EAL6	SML3 EAL6	SML3 EAL7

In examining at the corporation's security objectives, the customer and ISSE determined that *access control* to the sensitive data and *confidentiality* of the data as it transits over the intranet are the security services required. The mechanism(s) for implementation need to operate on both Windows NT and HP UNIX platforms.

The confidentiality mechanisms for the SML2 category recommend a minimum 80+ bit symmetric key length, 160+ exponent 1024+ modulus public key length. The firewall key scheme includes ISAKMP/OAKLEY with Data Encryption Standard (DES) or 3DES capability. 3DES is the scheme being evoked. The I&A mechanisms for the SML2 category recommend a system ID and a password with minimum character lengths. The corporation implements user IDs that are a minimum of six characters and eight characters for passwords with an alphanumeric mix. However, this is an internal Intranet, no security services for integrity, availability, and non-repudiation are considered necessary.

Each server requiring protection will have an individual firewall installed with the rules base requiring positive user identification and authentication before access is allowed. Initially, this process will be accomplished using use of user IDs and passwords; however, it will migrate to a PKI certificate based capability. Confidentiality will be provided by the Virtual Private Network (VPN) capability resident to the firewall product. Client VPN software will be installed on each client machine enforcing the connection and VPN rules to the protected servers (if the client VPN is disabled, no connection is allowed to a protected server).

The following security mechanisms are employed.

- Fronting each server that contains protected data with a firewall.
- Invoking VPNs between client machines and the server and printers (using 3DES algorithm).



- Implementing user identification and authentication using the VPN user ID and password.
- Implementing the firewall rule base to allow access by authorized only users from authorized workstations.
- Consideration is being given to replacing the VPN only client with a client that provides the VPN capability and extends the firewall policies to the user's desktop.

## 4.5.6 Robustness Strategy Evolution

Although “robustness” is now an inherent part of the IATF, it is a relatively new term in the IA lexicon and is not clearly seen as a unifying successor to a variety of similar existing concepts such as completeness, assurance, and accreditation.

The security mechanism tables provide guidance at three strength levels to support a variety of security services. At another level of table refinement, security functions would appear, each of which would implement a particular mechanism. For example, each specific cryptographic algorithm would be a security function to implement a cryptographic algorithm mechanism in support of, say, a confidentiality security service. Many security functions that implement each mechanism.

To compare and contrast these functions, there needs to be a way to cost the relative strengths. This effort would require developing cost metrics for each security service. Although functional specifications might be a relatively modest enhancement, the development of multiple costing schemes is likely to be a monumental effort. This level of refinement, which would enable uniform comparison of the protection provided by security mechanisms, is the goal of the strategy.

The IATF layered approach to security means that a variety of services and mechanisms may be necessary to achieve the necessary protection. A broader view needs to be developed, looking across all needed services and the mechanisms proposed to provide those services. The residual risk to a system product needs to be addressed based on the environment in which it is implemented.

In addition to the above concerns, and because threat environments and security technologies are changing continually, the guidance provided is subject to frequent revision. To the extent possible, all mechanism recommendations should be by indirect references to formally endorsed documents. In cases where this is not possible, periodic revision and trained ISSE application is the best way to ensure that guidance is current.

## 4.5.7 Real World Applications

In the real world, it quickly becomes too complicated and impractical to determine layered solution approaches and describe, offer, support, and implement them for more than a small number of different robustness levels. The threat levels and information value levels described earlier simply yield too many combinations of SML and EAL levels, as shown in Table 4-7. The Office of Secretary of Defense (OSD) Information Assurance guidance and policy for the DoD's Global Information Grid (GIG) divides robustness into three levels, a more practical approach.

The OSD GIG policy uses an implementation approach for robustness that makes conclusions based on real-world conditions (see Appendix E, OSD IA Policy Robustness Levels).

### 4.5.7.1 Future Work

The following areas need further attention.

- The Network Rating Model/Methodology also addresses “goodness.” How can that effort be incorporated into the strategy?
- The issue of composition of metrics needs to be addressed in the framework of layered security.
- There is a need to ensure that terminology used in the strategy is definitive and consistent with the remainder of the IATF.
- The current approach to security is considered non-scalable which means the process used for small systems may not be appropriate for large systems. This is also known as the composability problem and the layering problem. How can the robustness strategy help address this issue?
- The mechanism tables need to be reviewed for non-uniform level of detail and non-quantifiable entries.
- Update the strategy to incorporate Common Criteria “language” throughout, rather than just describing the EALs.
- Consider the effect of recommended robustness on return on investment to the customer.

## 4.6 Interoperability Framework

Users continue to become more dependent on information systems, fostering a driving need for connectivity and interoperability at the application level. As information and telecommunication systems are being introduced and updated, interoperability of these systems is a major concern of the organizations that use them. When these systems have to be secure, efficient interoperability becomes more difficult to achieve and manage. This section of the Framework provides a high level strategy for dealing with interoperability at the architecture and technology levels. Later releases of this Framework will address the issue of interoperability comprehensively, making users aware of options and tradeoffs, and providing guidance addressing with this important challenge.

### 4.6.1 Major Elements of Interoperability

This section identifies numerous elements that must be addressed in order to achieve interoperability. To achieve interoperability, it is typically necessary to deal with all of them. These elements and the issues associated with them are discussed below.

- *Architecture.* A first step in achieving interoperability is an agreement on the nature of the security services, the type of security mechanism(s) to be used, and their allocation

to functional components (e.g., enclave boundary interfaces, end user terminals of the architecture, and the layers at which security mechanisms are applied).

- *Security Protocols.* It is fundamental that systems use compatible communications protocols to achieve user-to-user connectivity. When this connectivity must be secure, several security elements associated with security protocols also must be considered. These elements include security services, cryptographic algorithms (with modes and bit lengths), synchronization techniques, and key exchange techniques. If options are permitted, common provisions are also needed for algorithm selection and broader security option negotiation. Typically, security protocol designers deal with these elements.
- *Product Compliance to Standards.* Another element needed for interoperability stems from the assurance that products used to implement a network security solution actually comply with the standards they claim to support. There are a number of initiatives with the commercial sector and in government that will verify compliance, as discussed below.
- *Interoperable KMI/PKI Support.* The services and techniques used to provide KMI/PKI are another element needed to achieve interoperability. This includes key and certificate formats, token mechanisms, cross certification (to facilitate communication across KMI/PKI security domains), directory systems, and compromise recovery capabilities. These considerations are discussed further in Section 4.7 (Key Management Infrastructure/Public Key Infrastructure Considerations).
- *Security Policy Agreement.* Beyond all of the technical issues needing to be addressed to allow interoperability is a fundamental issue of organizational security policies that establish the ground rules for permitting interoperability. It is necessary that the network or system “owners” determine what minimum protection mechanisms and assurances (perhaps for particular types of data or destinations) are needed before they would be willing to allow users from other networks or systems to communicate or interact with users of their resources and information. Because this important topic is beyond the scope of this document, it is assumed in this Framework that organizations wishing to interoperate have resolved any incompatibilities in organizational security policy and that the only barriers are technical or economic.

## 4.6.2 Challenges for Interoperability

When formulating an IA solution, the following potential impediments tend to act as obstacles to achieving interoperability:

- Backward compatibility with legacy systems that do not use accepted standards, and lack negotiation mechanisms needed to interoperate with newer standards-based implementations (even if backward-compatible protocols and modes are available).
- Security solutions lagging the rapid pace of evolving information technologies, often making security an adjunct capability.
- Evolution of standards or lack of standards accepted by either the user community or commercial product marketplace.
- Defacto proprietary standards or closed systems.

- Lack of an accepted source of testing to verify that products implementing standards do so correctly and that sufficient options of the standards are implemented, so users can be assured that the resultant products are in actuality, interoperable.

The challenge is to recognize and accommodate these obstacles, yet still find a way to achieve the interoperability needed by our customers.

### 4.6.3 Interoperability Strategy

At this point in the IATF, it is appropriate to establish a basic, high level strategy for dealing with interoperability. This strategy can be characterized by specific efforts focused on the following.

- Fostering standards for secure applications and communications protection that are based on open architectures.
- Supporting security negotiation protocol standards that allow users to have varying policies and provide a vehicle to negotiate elements of interoperability.
- Developing a migration strategy from the interim solutions to open standards in environments where emerging technology dominates and users accept interim solutions that are not standards based.
- Defining initial interoperability standards, and influencing and migrating to a standards-based approach where gaps exist.

A major issue still remains. It is imperative to ensure that products and system components correctly implement these standards and options so interoperability is actually realized. A number of initiatives within the government and private sectors exist to address this issue. These include the following.

- *Automotive Network eXchange® (ANX)*—The automotive industry has recognized the importance of interoperability for the transport of trading partner electronic information. The ANX network service is positioning to provide automotive trading partners with a single, secure network for electronic commerce and data transfer—replacing the complex, redundant, and costly multiple connections that exist throughout the automotive supply chain.
- *International Computer Security Association (ICSA)*—The ICSA promotes the open exchange of information between security product developers and security service providers. ICSA acts as an independent, third party that offers a number of initiatives, including a product certification program. The ICSA certification develops criteria by which industry wide categories of products are tested. The ICSA certifies products on an annual basis and spot-checks for compliance throughout the year against the latest version of each product. Using this process, buyers of ICSA-certified products can be assured of getting the most secure products available at the time.
- *National Information Assurance Partnership (NIAP)*—The NIAP is a joint industry/government initiative, lead by the National Institute of Standards (NIST) and NSA to establish commercial testing laboratories where industry product providers can have security products tested to verify their performance against vendor claims. As with the

ICSA initiatives, a natural fallout of this testing will be assurance that users can be assured that those products advertising compliance with standards will indeed be interoperable.

These activities, and a number of others similar to them, will help product and system providers to deliver solutions that will support the interoperability needs of their broad customer base.

The interoperability strategy presented in this section is embodied throughout this IATF. In a later release of the IATF document, a more detailed treatment of specific issues impacting interoperability will be included in subsequent sections. Specifically, Chapters 5 through 9 (Defend the Network and Infrastructure, Defend the Enclave Boundary/External Connections, Defend the Computing Environment, Supporting Infrastructures, and Information Assurance for the Tactical Environment) of the IATF will include discussions of interoperability issues specific to each of the user requirement categories. These will include interoperability concerns or needs reflected in the captured requirements, the technology assessments (to identify the degree to which the available solutions deal with interoperability issues), and the recommendations (that deal with selection of architectures and protocols that achieve the needed interoperability). Chapter 8 (Supporting Infrastructures) will deal specifically with interoperability issues associated with KMI/PKI.

## 4.7 Key Management Infrastructure/ Public Key Infrastructure Considerations

A KMI/PKI capability is needed to support most technical security countermeasures. This section provides a high level discussion of the role of, and features associated with, a KMI/PKI. Detailed guidance for the architecture of KMI/PKI can be found in Chapter 8 (Supporting Infrastructures) of the Framework.

### 4.7.1 KMI/PKI Overview

The KMI/PKI process generates, distributes, and manages security credentials. It can be considered as a set of interrelated activities providing security services that are needed to enable the Framework's security solutions presented in Chapters 5, 6, 7, and 9 (Defend the Network and Infrastructure, Defend the Enclave Boundary/External Connections, Defend the Computing Environment, Supporting Infrastructures, and Information Assurance for the Tactical Environment). KMI/PKI is a unique user requirement category in the Framework because it does not directly satisfy a user's security requirements; rather, it facilitates the use of security building blocks that are needed by other security mechanisms.

Current KMI/PKI implementations consist of numerous stovepipe infrastructures that support different user solutions. These are run by various organizations, even though the end user may need support from several stovepipes for a single application. A critical aspect of a complete system approach to any network security solution needs to include a KMI/PKI architecture that provides effective and efficient operations while maintaining the requisite security features and assurances.

A KMI/PKI architecture is heavily dependent on the specific applications it supports. For example, a VPN provides an encrypted pipe between two enclaves. The KMI/PKI provides keys and certificates to the cryptographic devices that provide authentication and encryption used to establish and maintain the pipe. KMI/PKI could also provide additional services, including data recovery and a directory to provide access to users' public certificates.

A second area where KMI/PKI differs from other solutions in the Framework is that its security is distributed throughout a number of separate elements. These elements require extensive security (e.g., encryption, certificate management, compromise recovery), among themselves to protect the user's key or certificate. Because of the repercussions of a successful attack against the KMI/PKI, internal infrastructure security requirements are often more stringent than user services security. There are also unique requirements also exist on the infrastructure (e.g., policy management), and the level of assurance for the KMI/PKI services is often higher.

## 4.7.2 KMI/PKI Operational Services

Section 8.1 (KMI/PKI) addresses four operational services supported by the KMI/PKI. These services support different user applications and consequently employ different (but related) mechanisms and have unique security requirements. The first user service is symmetric key generation and distribution. This is still the primary key management mechanism within the classified community.

The second service, PKI, addresses both digital signature (for authentication and integrity) and key agreement with its associated certificate management. This is the primary key management mechanism within the commercial community.

The third service, directory service, is used to provide access to the public information required with PKI such as the public certificate, the related infrastructure certificates, and the compromised key information. Directory services can be provided either by a global set of distributed directories (e.g., X.509 Defense Message System [DMS] directories), or by an on-line repository at a single site. Although directories can be used for other things, they are normally very closely coupled with PKI.

The final service is managing the infrastructure itself. The distributed nature of the infrastructure places additional functional and procedural requirements on the KMI/PKI and the sensitivity of the application places additional security requirements on the KMI/PKI. The internal structure of the infrastructure varies with the application it supports.

## 4.7.3 KMI/PKI Processes

The KMI/PKI consists of a numerous processes that all have to work together correctly for a user security service to be truly secure. Each of these processes is necessary at some level in all KMI/PKI architectures. The processes include the following.

- *Registration*—Authorizing people-making decisions about the validity of user actions.
- *Ordering*—Requesting the KMI/PKI to provide a user either a key or a certificate.

- *Key Generation*—Generating of the symmetric or asymmetric key by an infrastructure element.
- *Certificate Generation*—Binding the user information and the asymmetric key into a certificate.
- *Distribution*—Providing the keys and certificates to the user in a secure, authenticated manner.
- *Accounting*—Tracking the location and status of keys and certificates.
- *Compromise Recovery*—Removing invalid keys and certificates from the system in an authenticated manner.
- *Re-key*—Periodically replacing keys and certificates in a secure, authenticated manner.
- *Destruction*—Destructing the Secret key when it is no longer valid.
- *Data Recovery*—Being able to recover encrypted information without direct access to the original key.
- *Administration*—Running the infrastructure.
- *Value-added PKI Processes*—Supporting optional value-added processes including archive, time-stamp, and notary services, (PKIs only).

The complete set of KMI/PKI processes are usually distributed to several elements performing independent tasks, requiring extensive coordination and security processing between elements. For most processes, numerous ways exist to implement the services based on the application supported, the security required, and the cost (e.g., money, people, and performance) the user is willing to pay. Each process contributes to the overall security of the KMI/PKI and has various forms of attacks and countermeasures.

## References

1. Jeff Humphrey and Bruce Gabrielson, Ph.D. “Phreakers, Trashers, and Hackers,” presented at AFSEA INFOSEC Engineering Course, 1995, Burke, VA  
<http://blackmagic.com/ses/bruceg/hackers.html>
2. Root Shell at <http://www.rootshell.com>
3. Coast Security Pages at <http://www.cs.purdue.edu/coast/intrusion-detection/>
4. FIPS PUB 140-1, National Institute of Standards and Technology, Security Requirements for Cryptographic Modules, <http://www.itl.nist.gov/div897/pubs/fip140-1.htm>
5. NSTISSI No. 4009, National INFOSEC Glossary.
6. Common Criteria for Information Technology Security Evaluation, CCIB-98 (ISO/IEC 15408), Version 2.0, 1998, <http://csrc.nist.gov/cc/>.
7. DoD Reg. 5200.1-R, Information Security Program, 1997.
8. NSTISSAM TEMPEST/1-92, Compromising Emanations Laboratory Test Requirements, Electromagnetics, 1992.
9. Laing, Alan, “DoD PKI Level of Protection and The Appropriateness of Proposed Solutions for Various Applications”, Draft ,1998.
10. National Security Agency Specification for General Functional Security Requirements for a Telecommunications System (FSRS), 1989.
11. Information Systems Security Engineering Handbook, Release 1.0, 28 February 1994.
12. Security Management Infrastructure (SMI) Task 1 Team, Threat and Vulnerability Model for Information Security, 1997.

## Additional References

- a. NSA/CSS Dir. No. 120-1, NSA/CSS Operations Security Program, 1990.
- b. National Security Agency Specification for Unified INFOSEC Criteria, 1991.
- c. 16 Warwick Ford, *Computer Communications Security*, Prentice Hall PTR, Englewood Cliffs, NJ, 07632, 1994.



## Chapter 4

# Technical Security Countermeasures

The authors of the Information Assurance Technical Framework IATF recognize the importance of using both technical and non-technical countermeasures in formulating an effective overall security solution to address threats at all layers of the information infrastructure. This chapter of the IATF presents a general discussion of the principles of determining appropriate technical security countermeasures. It includes a detailed assessment on threats of, important security services, robustness strategy, interoperability framework, and the Key Management Infrastructure (KMI)/Public Key Infrastructure (PKI). It also provides background for the detailed technical discussions contained in later sections of the IATF.

## 4.1 Introduction

Adversaries' primary goals can be grouped into three general categories: unauthorized access, unauthorized modification, and denial of authorized access. The intent of a security solution is to prevent an adversary from successfully achieving these goals. This chapter discusses threats, security services and appropriate security technologies. Using the methodology described in Chapter 3 (Information Systems Security Methodology), with the consideration of applicable threats, security solutions may be proposed which support appropriate security services and objectives. Subsequently, proposed security solutions may be evaluated to determine if residual vulnerabilities exist, and a managed approach to mitigating risks may be proposed. *Security services* are those services that safeguard and secure information and information systems. Access Control, confidentiality, integrity, availability, and non-repudiation are the five primary security services. These services are accomplished by incorporating security mechanisms, e.g., encryption, identification, authentication, access control, security management and trust technology into the information system to form a barrier to attack. This chapter presents an overview (including a definition) of each of these services, a breakdown of the various elements included in each and a detailed look at the security mechanisms that support each service.

Three additional topics, robustness, interoperability, and KMI/PKI should be considered in selection of security countermeasures. The *robustness strategy* provides a philosophy and initial guidance for selecting the strength of security mechanisms and the security assurance provisions that may be needed for a particular value of information and a potential threat level. This section defines the IATF strategy for measuring and assessing the need for various levels of robustness for technical (and selected non-technical) security countermeasures. The robustness strategy is not intended to provide universal answers on needed strength or assurance that is, it is not a "cookbook." The final selection of mechanisms, and the necessary level of strength and assurance needed will be based on an Information Systems Security Engineering (ISSE)

activity and a resultant risk management process that addresses the situation of a specific user, mission, and environment.

The robustness of a security solution must be considered in relation to the system requirement for connectivity. Recognizing the growing need for connectivity, an *interoperability framework* provides a strategy to ensure that security provisions (1) do not inhibit the connectivity that is available without security and (2) if needed, maintain backward compatibility with existing system capabilities. The chapter continues with a discussion of *KMI/PKI Considerations*. Typically, the need for KMI/PKI capabilities accompanies the use of technical security countermeasures. It is important to consider the needs that a KMI/PKI creates and the demands it places on network users and operators in the context of any potential network security solution.

This chapter provides a basic framework for considering these important topics. Each facet of a solution addressed in this chapter should be considered in relating to the other facets. For example, the robustness of a solution depends on the way the technology is implemented. Similarly, knowledge of the primary security services and the important security technologies will facilitate the formation of effective security solutions. In addition, considering interoperability and KMI/PKI concurrently with the formulation of a security solution will help to ensure the effectiveness of that solution.

## 4.2 Adversaries, Motivations, Capabilities, and Categories of Attacks

Adversaries come from various backgrounds and have a wide range of financial resources at their disposal. In this section a host of potential adversaries are examined. What produces an adversary? What are each adversary's motivations? What category(s) of attacks does each adversary use? This section seeks to answer these questions by providing information on the various potential adversaries and by providing examples of attacks in each attack category along with a brief description of how each attack is performed and by whom.

This section also discusses the countermeasures that can be used against potential adversaries and the different categories of attack.

### 4.2.1 Potential Adversaries

One typically thinks of adversaries as having malicious intent. However, in the context of system security and protecting one's systems and information, it is also important to consider the threat posed by those without malicious intent. Table 4-1 provides examples of individuals and organizations in both of these categories.

**Table 4-1 Examples of Potential Adversaries**

Adversary	Description
<b>Malicious</b>	
Nation States	State-run, well organized and financed. Use foreign service agents to gather classified or critical information from countries viewed as hostile or as having an economic, military or a political advantage.
Hackers	A group of individuals (e.g., hackers, phreakers, crackers, trashers, and pirates) who attack networks and systems seeking to exploit the vulnerabilities in operating systems or other flaws.
Terrorists/ Cyberterrorists	Individuals or groups operating domestically or internationally who represent various terrorist or extremist groups that use violence or the threat of violence to incite fear with the intention of coercing or intimidating governments or societies into succumbing to their demands.
Organized Crime	Coordinated criminal activities including gambling, racketeering, narcotics trafficking, and many others. An organized and well-financed criminal organization.
Other Criminal Elements	Another facet of the criminal community, which is normally not very well organized or financed. Normally consists of very few individuals, or of one individual acting alone.
International Press	Organizations that gather and distribute news, at times illegally, selling their services to both print and entertainment media. Involved in gathering information on everything and anyone at any given time.
Industrial Competitors	Foreign and domestic corporations operating in a competitive market and often engaged in the illegal gathering of information from competitors or foreign governments in the form of corporate espionage.
Disgruntled Employees	Angry, dissatisfied individuals with the potential to inflict harm on the local network or system. Can represent an insider threat depending on the current state of the individual's employment and access to the system.
<b>Non-Malicious</b>	
Careless or Poorly Trained Employees	Those users who, either through lack of training, lack of concern, or lack of attentiveness pose a threat to information systems. This is another example of an insider threat or adversary.

### 4.2.1.1 Motivations

Individual motivations to “get inside” are many and varied. Those with malicious intent wishing to achieve commercial, military, or personal gain are known as “hackers”[1]. At the opposite end of the spectrum are those who accidentally do something that compromises the network. Hackers range from the inexperienced professional, college student, or novice (e.g., Script Kiddy) to the highly technical and very capable (e.g., Uberhacker). Most hackers pride themselves on their skill and seek, not to destroy, but simply to gain access so that the computer or network can be used for later experimentation. Hackers often believe that by exposing a hole or “back-door” in a computer system, they are actually helping the organization to close the holes, providing an actual benefit to the Internet and a needed resource. Other hackers have less benign motives for “getting inside”.

Intelligence gathering, information operations, or psychological warfare are some motivations behind attempts to gain access. The following are some common reasons why an adversary might be motivated to exploit a particular target.

- Gain access to classified or sensitive information. (Note: What is of high value to one person or organization may be of no value to another).
- Track or monitor the target's operations (traffic analysis).
- Disrupt the target's operations.
- Steal money, products, or services.
- Obtain free use of resources (e.g., computing resources or free use of networks).
- Embarrass the target.
- Overcome the technical challenge of defeating security mechanisms.

From an information system's standpoint, these motivations can express themselves in three basic goals: access to information, modification of or destruction of information or system processes, or denial of access to information. In attacking an information processing system an adversary accepts a certain amount of risk. This risk may be time dependent. The risk of loss to the adversary may far exceed the expected gain. Risk factors include:

- Revealing the adversary's ability to perform other types of attacks,
- Triggering responses that might prevent the success of a future attack especially when the gain is much greater,
- Incurring penalties, (e.g., fines, imprisonment, embarrassment), and
- Endangering human life.

The level of risk that an adversary is willing to accept depends on the adversary's motivation.

## 4.2.1.2 Capabilities

Adversaries' capabilities determine their ability to implement attacks against the information processing system. Some capability factors are:

- Knowledge and skills in developing attacks,
- Availability of necessary resources.

The greater the capabilities of the adversary, the greater the likelihood of an attack. If the adversary has the necessary knowledge, skills, and resources and is willing to put themselves and those resources at risk, then the only remaining factor is opportunity. Although opportunity is not captured in our list of capabilities it is the last key element that is necessary if an adversary is to attack. Opportunity may present itself in a number of forms including vulnerabilities in a particular operating system, misconfiguration of routers or firewalls, and modems attached but unprotected inside the local enclave. Reducing an adversary's capabilities usually is not possible but reducing the adversary's opportunity is.

## 4.2.2 Categories of Attacks

Chapter 1 (Introduction), Table 1-1, addresses five categories of system attack. As shown in Figure 4-1, each of these has unique characteristics that should be considered in defining and implementing countermeasures. This section provides an overview of each category of attack,

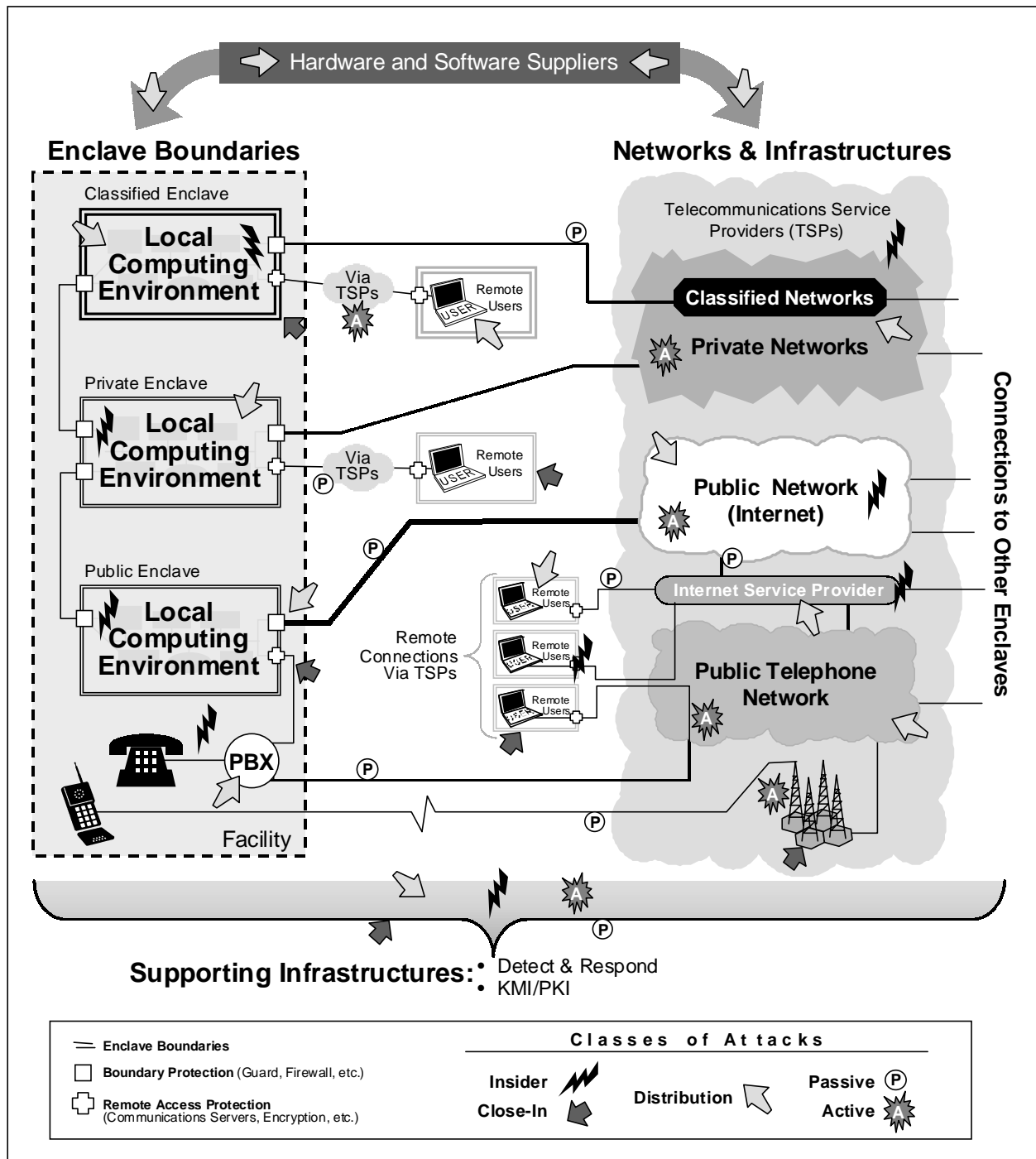


Figure 4-1. Categories of Attacks Against Networked Systems

with specific examples of attacks for each category of attack. Note that all network-based attacks are combined in the following discussion.

### 4.2.2.1 Passive Attacks

These attacks involve passive monitoring of communications sent over public media (e.g., radio, satellite, microwave, and public switched networks). Examples of countermeasures against these attacks include the use of VPNs, cryptographically protected networks, and use of protected distribution networks (e.g. physically protected/alarmed wire-line distribution network). Table 4-2 provides examples of specific attacks characteristic of this class.

**Table 4-2. Examples of Specific Passive Attacks**

Attack	Description
Monitoring Plaintext	An attacker monitoring the network could capture user or enclave data that is not otherwise protected from disclosure.
Decrypting Weakly Encrypted Traffic	Crypto-analytic capability is available in the public domain, as witnessed by the June 1997 collaborative breaking of the 56-bit-strength Data Encryption Standard (DES). While the near-term threat to large volumes of traffic is questionable given the number of machines and hours involved, it does show the vulnerability of any single transaction.
Password Sniffing	This type of attack involves use of protocol analyzers to capture passwords for unauthorized reuse.
Traffic Analysis	Observation of external traffic patterns can give critical information to adversaries even without decryption of the underlying information. For instance, extension of a network into a tactical theater of operations. Changes in traffic patterns may indicate the imminence of offensive operations thereby removing the element of surprise.

### 4.2.2.2 Active Attacks

Active attacks include attempts to circumvent or break security features, introduce malicious code (such as computer viruses), and subvert data or system integrity. Typical countermeasures include strong enclave boundary protection (e.g., firewalls and guards), access control based on authenticated identities for network management interactions, protected remote access, quality security administration, automated virus detection tools, audit, and intrusion detection. Table 4-3 provides examples of specific attacks characteristic of this class.

**Table 4-3 Examples of Active Attacks**

Attack	Description
Modification of Data in Transit	In the financial community, it could be disastrous if electronic transactions could be modified to change the amount of the transaction or redirect the transaction to another account.
Replay (Insertion of Data)	Re-insertion of previous messages could delay timely actions. Bellovin shows how the ability to splice messages together can be used to change information in transit and produce desired results.
Session Hijacking	This attack involves unauthorized use of an established communications session.

Attack	Description
Masquerade as Authorized User/Server	This attack involves an attacker identifying himself or herself as someone else, and thereby gaining unauthorized access to resources and information. An attacker gets user/administrator information by employing sniffers or other means, then uses that information to log in as an authorized user. This class of attack also includes rogue servers that can be used to obtain sensitive information after establishing what is believed to be a trusted service relationship with the unsuspecting user.
Exploit System-Application and Operating System Software	An attacker exploits vulnerabilities in software that runs with system privileges. Well-known attacks involve sendmail and X-Windows server vulnerabilities. Recently, there has been an increase in alerts regarding Windows 95 and Windows NT vulnerabilities. New vulnerabilities for various software and hardware platforms are discovered almost daily. Attacks, vulnerabilities, and patches are reported through the various computer emergency response alerts and bulletins.
Exploiting Host or Network Trust	An attacker exploits transitive trust by manipulating files that facilitate the provision of services on virtual/remote machines. Well-known attacks involve rhosts and .rlogin, which facilitate workstations sharing of files and services across an enterprise network.
Exploiting Data Execution	An attacker can get the user to execute malicious code by including the code in seemingly innocent software or e-mail that is downloaded. The malicious code might be used to destroy or modify files, especially files that contain privilege parameters or values. Well-known attacks have involved PostScript, Active-X, and MS Word macro viruses.
Inserting and Exploiting Malicious Code (Trojan Horse, Trap Door, Virus, Worm)	An attacker can gain execution access to a user's system commands through one of the vulnerabilities previously identified and use that access to accomplish the his or her objectives. This could include implanting software to be executed based on the occurrence of some future event. Hacker tools such as Rootkit (see <a href="http://www.rootshell.com">http://www.rootshell.com</a> [2] to download Rootkit or any of a large number of security/hacker tools) have turnkey capabilities, including an insertion script, root grabbing, Ethernet sniffing, and track hiding to mask the presence of a hacker.
Exploiting Protocols or Infrastructure Bugs	<p>An attacker exploits weaknesses in protocols to spoof users or reroute traffic. Well-known attacks of this type include spoofing domain name servers to gain unauthorized remote login, and bombing using Internet Control Message Protocol (ICMP) to knock a machine off the air. Other well-known attacks include source routing to impersonate a trusted host source, Transmission Control Protocol (TCP) sequence guessing to gain access, and TCP splicing to hijack a legitimate connection.</p> <p>Malicious code can exfiltrate information through a lower level tunnel within a virtual private network. At least one published paper points out potential security concerns revolving around use of Internet Protocol Security (IPSec) default security mechanisms. What Bellovin points out are occasions on which the integrity functions of DES in Cipher Block Chaining mode can be circumvented with the right applications by splicing of packets.</p>
Denial of Service	An attacker has many alternatives in this category, including ICMP bombs to effectively get a router off the network, flooding the network with garbage packets, and flooding mail hubs with junk mail.

### 4.2.2.3 Close-in Attacks

Close-in attacks are attacks in which an unauthorized individual gains close physical proximity to networks, systems, or facilities for the purpose of modifying, gathering, or denying access to, information. Gaining such proximity is accomplished through surreptitious entry, open access, or both. Table 4-4 provides examples of specific attacks characteristic of this class.

**Table 4-4 Examples of Close-In Attacks**

Attack	Description
Modification of Data/Information Gathering	This results from an individual in close proximity gaining physical access to the local system and, as a result, modifying or stealing information such as IP addresses, login ID schemes, and passwords.
System Tampering	This type of attack results from an individual in close proximity gaining access to and tampering with the system (e.g., bugging, degrading).
Physical Destruction	This type of attack results from an individual in close proximity gaining physical access, resulting in the physical destruction of a local system.

### 4.2.2.4 Insider Attacks

Insider attacks are performed by a person who either is authorized to be within the physical boundaries of the information security processing system or has direct access to the information security processing system. There are two types of insider attacks: malicious and non-malicious (carelessness or ignorance of the user). The non-malicious case is considered an attack because of the security consequences of the user's action.

- *Malicious Insider Attacks*—The Federal Bureau of Investigation (FBI) estimates indicate that 80 percent of attacks and intrusions come from within organizations (see <http://www.cs.purdue.edu/coast/intrusion-detection/>) [3]. An insider knows the layout of the system, where the valuable data is, and what security precautions are in place. Insider attacks originate from within the enclave and are often the most difficult to detect and to defend against.

Sources of insider attacks can include uncleared cleaning crews (with after-hours physical access), authorized (privileged to login) system users, and system administrators with malicious intent. Often it is difficult to prevent individuals who have legitimate access to a system from transgressing into more private areas where they have no authorization. Insider attacks may focus on compromise of data or access and can include modification of system protection measures. A malicious insider may use covert channels to signal private information outside of an otherwise protected network. However, there are many other avenues through which a malicious insider can damage an information system.

- *Non-malicious Insider Attacks*—These attacks are caused by authorized persons who have no intent to cause damage to the information or to the information processing system but may unintentionally do so by some specific action. The damage may be caused by lack of knowledge or by carelessness.



Typical countermeasures include security awareness and training; audit and intrusion detection; security policy and enforcement; specialized access control of critical data, servers, Local Area Networks (LAN), etc. implemented by trust technology in computer and network elements; or a strong Identification and Authentication (I&A) capability. Table 4-5 contains examples of specific attacks characteristic of this class.

**Table 4-5 Examples of Insider Attacks**

<b>Attack</b>	<b>Description</b>
<b>Malicious</b>	
Modification of Data or Security Mechanisms	Insiders often have access to information due to commonality of shared networks. This can give the insider access, allowing manipulation or destruction of information without authorization.
Establishment of Unauthorized Network Connections	This results when users with physical access to a classified network create an unauthorized connection to a lower classification level or sensitivity network. Typically this is in direct violation of the classified network's security policy or user directives and procedures.
Covert Channels	Covert channels are unauthorized communication paths used for transferring misappropriated information from the local enclave to a remote site.
Physical Damage/ Destruction	This is intentional damage to or destruction of a local system(s) resulting from the physical access afforded the insider.
<b>Non-Malicious</b>	
Modification of Data	This type of attack results when insiders either through lack of training, lack of concern, or lack of attentiveness, modify or destroy information located on the system.
Physical Damage/ Destruction	This type of attack is listed under malicious as well. As a non-malicious attack, it can result from carelessness on the part of the insider, for instance, failure to obey posted guidance and regulations resulting in accidental damage to or destruction of a system.

## 4.2.2.5 Distribution Attacks

The term “distribution attack” refers to the potential for malicious modification of hardware or software between the time of its production by a developer and its installation or when it is in transit from one site to another. The threat at the factory can be minimized by strong in-process configuration control. Distribution threats can be addressed by use of controlled distribution, or by signed software and access control that is verified at the final user site. Table 4-6 contains examples of specific attacks characteristic of this class.

**Table 4-6 Examples of Distribution Attacks**

Attack	Description
Modification of Software/Hardware at Manufacturer's Facility	These attacks can be performed by modifying the configuration of software or hardware while it is cycling through the production process. The countermeasures for threats during this phase include rigid integrity controls including high-assurance configuration control, and cryptographic signatures on tested software products.
Modification of Software/Hardware during Distribution	These attacks can be performed by modifying the configuration of software or hardware during its distribution (e.g. embedment of listening devices during shipment). The countermeasures for threats during this phase include use of tamper detection technologies during packaging, use of authorized couriers and approved carriers, and use of blind-buy techniques.

## 4.3 Primary Security Services

The IATF guidance incorporates five primary security services: access control, confidentiality, integrity, availability, and non-repudiation. The division of network security principles into standard security service categories is convenient for this description. The categories presented below roughly follow the “basic security services” identified in the 1990 Recommendation X.800, “Security Architecture for Open Systems Interconnection for Consultative Committee for International Telephone and Telegraph (CCITT) Applications” (which is technically aligned with ISO 7498-2, “Information Processing Systems Open Systems Interconnection, Basic Reference Model”, Part 2: Security Architecture) and more recently, the International Organization for Standardization (ISO)/International Engineering Consortium (IEC) 10181 series, Parts 1-7.

In practice, not one of these security services is isolated or independent of the other services. Each service interacts with and depends on the others. For example, confidentiality is of limited value unless preceded by some type of authorization process. Protection from unauthorized entities is not possible if one cannot determine whether the entity he or she is communicating with is authorized. In actual implementations, lines between the security services are blurred by the use of specific mechanisms that contribute to supporting more than one service.

Given this caveat, this section characterizes each service according to its basic functional elements and discusses the mechanisms that are available to implement the elements of that service. Where appropriate, considerations of the relative strengths of these mechanisms are also noted.

### 4.3.1 Access Control

In the context of network security, access control means limiting access to networked resources (hardware and software) and data (stored and communicated). The goal of access control is to prevent the unauthorized use of these resources and the unauthorized disclosure or modification of data. Access control also includes “resource control”, for example, preventing logon to local workstation equipment or limiting use of dial-in modems. For the purposes of this discussion, network access control is not concerned with denying physical access (e.g., via locked rooms or tamperproof equipment).

Access control is applied to an “entity” based on an Identity and/or an Authorization. An identity may represent an actual user, a process with its own identity (e.g., a program making a remote access connection), or a number of users represented by single identity (e.g., role based access control).

Access control mechanisms are most often used as a set of mechanisms, which may be used by other security services. Confidentiality, integrity, availability, and limiting use of network resources all depend on limiting the ability of an adversary to access an item or service.

The elements of access control can be categorized in the following areas.

- *I&A* – Establishing the identities of entities with some level of assurance (an authenticated identity).
- *Authorization* – Determining the access rights of an entity, also with some level of assurance.
- *Decision* – Comparing the rights (authorization) of an authenticated identity with the characteristics of a requested action to determine whether the request should be granted.
- *Enforcement* – Enforcement may involve a single decision to grant or deny or may entail periodic or continuous enforcement functions (continuous authentication).

The following subsections discuss these elements and provide examples of the mechanisms that are available to implement them.

### 4.3.1.1 I&A

I&A is a set of security services used in conjunction with most other security services. The first step of most security services is to determine the identities of one or more of the parties participating in an action. A trusted identity must be used for access control decisions as well as to provide non-repudiation and accountability evidence. Knowing the identity of an entity and the existence of a peer relationship is also fundamental to establishing communication with confidentiality and integrity. If the identity of the peer in a secure communications path is not properly established, it leaves open the possibility that an unauthorized user (an adversary) could masquerade as an authorized user, leaving the data open to disclosure or manipulation by the adversary.

The process of determining an authentic identity consists of the following.

#### 4.3.1.1.1 Assigning, Binding, and Representing

There must be a mechanism for providing some assurance in the assignment of an identity. The entity that assigns identity must have a position with some level of trust (either implied or assured by a third entity common to both with a higher position or level of trust. These trusted entities must implement a process of identity checking that protects against assignment of improper identities. Process examples include checking driver’s licenses or verifying fingerprints. Assigning identity is the equivalent of a registration process and could take place through an existing security mechanism with its own identity establishment mechanism.

An identity must be unique amongst the community that will be validating that identity. This requires implementation of a community wide assignment mechanism that provides a unique identity to each entity. An identity also must be conveyed in a representation that is recognized and can be commonly processed by all potential peers to a communication. This implies a standard format for representing identity.

Identities used for network access control can be assigned and represented by many different mechanisms; for example:

- System administrators providing accounts and passwords for UNIX usernames,
- Network administrators assigning Internet Protocol (IP) addresses to machines,
- Key distribution methods that distribute symmetric keys,
- Key distribution methods that distribute public/private key pairs,
- Certification Authorities (CA) generating public key certificates containing Distinguished Names (DN), and
- Security officers associating a set of fingerprints with a common name.

The assurance level attributed to an identity depends on the processes used to verify the correctness of that identity before it is issued, the trust instilled by the entity assigning the identity, and the strength of the binding between the entity and the identity. Verification may range from requesting a mother's maiden name over the telephone to checking driver's licenses or verifying fingerprints in person. Means of instilling trust in issuers include procedural mechanisms, such as a company's assigning system administrators, legal mechanisms, such as notaries, and technological mechanisms such as certification paths in a certification hierarchy. Mechanisms for binding entities to identities include signed X.509 certificates and password files associated with access control lists.

Strongly establishing identities for communicating entities is the first step in countering any attack that is predicated on adversaries, representing themselves as someone or something that they are not (including masquerading and insider modification attacks).

#### 4.3.1.1.2 Communicating and Authenticating

To authenticate an entity that is attempting to gain access, an identity must be associated with the access request and provided to the communicating peer. Along with an indication of identity, the authenticating peer must have the parameters (authentication information) needed to validate that identity. Authentication is implemented by user-to-host and peer-to-peer, and Trusted Third Party (TTP) architectures as follows.

- User-to-Host—When a user logs onto a host (or workstation), the user must be identified and authenticated before access to the host or network is granted. This process requires a mechanism to authenticate a real person to a machine. The best methods of doing this involve multiple forms of authentication such as password, physical token, and biometric verification (e.g. something you *know*, something you *have*, something you *are*).
- Peer-to-Peer Authentication—A peer-to-peer authentication architecture, sometimes referred to as mutual authentication protocol, involves the direct communication of

authentication information between the communicating entities (e.g., peer-to-peer or client host-to-server). No other entities are required. This architecture is possible only if each entity in a security domain is able to obtain the authentication information of every communicating entity in the domain.

- **Trusted Third Party Authentication**—The architecture for TTP authentication uses a third entity, trusted by all entities, to provide authentication information. A TTP may provide authentication information in each instance of authentication, in real-time, or as a precursor to an exchange (such as a certificate authority). The amount of trust given the third party must be evaluated. Methods of establishing and maintaining a level of trust in a TTP include certification practice statements that establish rules, processes, and procedures that a CA uses to assure the integrity of the authentication process and use of secure protocols to interface with authentication servers.

The mechanisms used for authenticating of an identity can be categorized as simple or cryptographically based. Simple mechanisms may include identification based on identities which are verified by asking the entity to communicate information that only the entity attempting access would know (e.g. a password and locally stored password file). Assurance comes from the local binding between the password and an identity. Another example of a simple authentication method is address-based authentication. Address-based mechanisms authenticate identity based solely on assigned network addresses (e.g., Internet Protocol (IP) address) of communicating peers as compared to known IP address assignment for the entities.

Cryptographic-based mechanisms rely on the cryptographic processing of data within a defined protocol. Peers may share a common secret key (often stored in a hardware token) to process, or encrypt the exchange in a challenge-response protocol. Other cryptographic mechanisms rely on public key cryptography alone, or on the binding between a public key and an identity provided by public key certificates. Examples of how an identity is authenticated in each cryptographic technique are provided below.

- **Identity Is A Locally Defined Name**—Identities of all potential communicating peers are stored locally in a trusted database that associates identities with their public keys. These public keys correspond to the private key used to sign a unique piece of data. Verifying a signature by using a stored public key authenticates an identity.
- **Identity Means the Defined Name**—From the valid X.509 certificate containing the public key that corresponds to the private key used to sign a unique piece of data. A valid X.509 certificate means that the complete certification path has been validated (including Certificate Revocation List (CRL) and Compromised Key List (CKL) checks and validity periods for all certificates) to a trusted root. X.509 certificates (of communicating peers or of the entities in certification paths) may be stored locally (cached), carried in the security association protocol, or accessed as needed from an X.500 directory, or any combination of these three methods. Verifying a signature by using a valid public key authenticates an identity.

For all cryptographically based mechanisms, the strength of the mechanism lies partly in the strength of the cryptographic algorithms (including key size), partly in the security of any communications protocol, and in large part, in the protection provided to secret key material.

There are a number of mechanisms for implementing and distributing identity and authentication information. Some of these mechanisms are:

- Names and passwords stored in a database local to the entity making the access control decision.
- IP addresses provided by a secure Domain Name Server (DNS).
- Passwords generated locally based on time (one time passwords).
- Symmetric keys stored in a local database.
- Public keys stored in a local database,
- Public key certificates provided by directories in response to queries.
- Authentication information may be carried in the communications protocols themselves.

The assurance level of the communication of identity and authentication information processes depends on whether that information needs protecting and how well it is protected. For example, passwords are sensitive because they can be used by anyone who knows them; they should therefore be encrypted for storage and transport. Certificates can be stored in unprotected directories or carried on unencrypted communications channels because they can only be used by the entity that holds the associated private key.

Note that identity information and the information used to authenticate that identity do not have to flow over the same communications path. A common example is name and password logins. Users are queried for a name and an associated password (the identity information) over the communications protocol. The authenticity of that name and password pair is established only by checking a locally stored database (the information used to authenticate provided by an off-line process).

There are entire infrastructures devoted to providing identities and the means of authenticating those identities. Examples of infrastructures supporting the determination of an authentic identity include the X.509 authentication framework, the Internet Engineering Task Force (IETF) public key infrastructure, the secure DNS initiatives, and the Simple Public Key Infrastructure (SPKI).

### 4.3.1.2 Authorization

Another important step in an access decision is determining the authorizations of one or more of the parties participating in a communication. These authorizations result in the granting of a set of privileges to an entity. Much like identity, authorizations must be conveyed in a commonly understood format and must be presented or maintained with some level of confidence. The process of determining an authenticated set of authorizations generally consist of the same components as determining an authenticated identity. A strong mechanism for determining authorizations can prevent an attack in which an entity attempts to forge access rights from being successful.

The process of determining the authorizations of an entity consists of assigning authorizations, binding authorizations to an entity, representing those authorizations in a standard format,

communicating those authorizations, and establishing the authenticity of the authorizations. These steps are discussed below.

#### 4.3.1.2.1 Assigning, Binding, and Representing

As in assigning identity, the process that determines and assigns authorizations must evoke a level of trust. Responsibility for that process falls on roles with names like certificate authority (CA), attribute authority, Access Control List (ACL) administrator, and system administrator. Authorizations used for network access control can be assigned by:

- System administrators who assign usernames to groups,
- Data owners who grant authorizations to read/write/execute files,
- Network administrators who generate ACLs,
- X.500 CAs who generate version 3 X.509 certificates containing extensions, and
- Attribute authorities who generate ANSI X9.57 attribute certificates.

#### 4.3.1.2.2 Communicating and Authenticating

Communicating authorization information follows the same model as authentication information. It may be pre-distributed and stored at each entity (e.g., ACLs); it may be carried in the communications protocol; or it may be provided by a trusted third party (e.g., X.500 directory, Radius authentication servers). There are a number of models for distributing authorization information:

- ACLs stored local to the entity making the access control decision.
- X.500 directories deployed to provide X.509 certificates.
- X.500 directories deployed to provide attribute certificates.
- Authenticity of authorization information is provided either by its trusted relationship with identity information (local binding) or because it is carried in cryptographically verifiable certificates.

The level of trust attributed to the third parties used for obtaining authorization information (either the parties who generated authorizations initially or those that distribute them when needed) is always an issue. The cryptographic techniques invoked to prove the authenticity of X.509 certificates and to bind attribute certificates to identity certificates represent one attempt to assure that trust.

### 4.3.1.3 Access Control Decision

The components discussed previously provide the information required to make an access control decision. They provide mechanisms for determining both the identity and the privilege set of a communicating entity. In practice, access decisions are usually based on an access control policy commonly referred to in the classified arena as discretionary or mandatory policies. International standards do not use the “mandatory/discretionary” terminology but rather Identity Based Access Control (IBAC), which bases decisions on an identity or Rule Based Access Control (RBAC), which checks an entity’s authorizations against an established rule set.

Within the scope of this discussion, IBAC and discretionary policies can be considered equivalent, and RBAC and mandatory policies can be considered equivalent. In either case, the function of an access control decision is to result in a simple grant or deny response to requests for access.

An IBAC decision grants or denies a request based on the presence of an entity on an ACL. If an entity is on the ACL, access to the requested information or resource is permitted—otherwise, access is denied. IBAC requires an authenticated identity prior to granting any access.

An RBAC decision depends on policies that can be algorithmically expressed and thus implemented on a computer system. These policies are stated such that resources have restrictions and entities must have authorizations. Access is granted to a resource based on an entity's authorizations rather than an entity's identity. An RBAC decision requires authorization information and restriction information to compare before any access is granted.

A composite policy, referred to as role-based policy, can be considered a variation of both IBAC and RBAC. An identity is assigned to a group that has been granted authorizations. Identities can be members of one or more groups. A current example is the Global Command and Control System (GCCS), which depends on organizational and role associations.

Most network operating systems have their own method of implementing access control, but they are all identity-based IBAC. Entities are granted access to resources based on an identity established during network login, which is compared with one or more ACL. These lists may be individually administered, may be centrally administered and distributed to individual locations, or may reside on a central server(s).

Mechanisms for establishing identities and authorizations are discussed in previous sections. Mechanisms for establishing the restrictions on a resource must be provided to implement an RBAC scheme. Since rule-based access controls how rules are implemented primarily in systems dealing with sensitive information, restrictions are most often expressed as policies for accessing sensitive data. To facilitate these policies, the sensitivities of a data item are conveyed in a data label and must be compared with the set of privileges assigned to an entity. Access is granted to sensitive information if an entity's privileges are appropriate for the sensitivities of the data. An example of a rule-based policy is the classifications used to distinguish information on a national security level, such as Top Secret, Secret, and Confidential, and the rule that identities granted authorization for any security level are authorized access to all lower security levels. Users who hold Secret clearances will be allowed to access Secret and below classified information.

Consistent with the issues surrounding identities and authorizations, data labels must also be assigned, bound, represented, communicated, and authenticated. There are currently many representations of a data security label (Federal Information Publications (FIPS)[4] 188 Standard Security Label, SDE Security Label—IEEE 802.10g, Internet Security Label, International Organization of Standardization (ISO) SC-27 Security Label, Common Security Label (Military Standard [MIL STD] 2045-48501), X.411 MHS: MTS Service Definition—Security Label). Establishment of a universally accepted standard is an area for further work.

Note that an access request can actually be composed of a complicated set of parameters. For example, a particular access might be- Execute a file labeled Top Secret at 3:15 p.m. during a



time of war. Defining “access” in this manner allows the access decision function to provide a binary “grant” or “deny” result. This introduces a new set of information that must be represented, communicated, and authenticated including contextual information, such as time, status, or current conditions.

### 4.3.1.4 Enforcement

Actual enforcement of the access control decision is really the step that provides the protection against threats. All previously discussed mechanisms for preventing attacks come together here with the enforcement of those protections.

The concept of enforcing an access control decision is separate from the decision itself. This is because the two processes may reside in different places architecturally. This separation permits the concept of an “authentication server” that makes an access decision for the network communications process to allow or prevent a requested access from taking place. For example, the access decision may result in the subject’s being provided with a token (such as a certificate) that guarantees the subject the right to access its target up to but no more than  $n$  times before a given time. This token is called a ticket or capability. These tokens may be cached at the target to improve efficiency.

An access control decision and its enforcement can be made at either end of a communications association. An example is the difference between a client accessing a File Transfer Protocol (FTP) server (the server limits access to files after a client request is submitted) and an electronic mail (e-mail) message (in which the originator decides whether the recipient should receive the message before a connection is made). In the e-mail example, the recipient’s mail software may also perform an additional access control check to determine whether the recipient can be allowed to view the message.

Another distinction between access control mechanisms is whether the decision and enforcement process occurs once at the initiation of a communications session, is repeated periodically throughout a session, or qualifies as “continuously authenticated.” A method commonly used to assure that access to a communications session is controlled continuously is use of encryption mechanisms to prevent loss of control of the session (session stealing or hijacking). Indeed, it can be argued that access is not completely controlled if information flowing over a public network is not protected by the confidentiality security service.

Enforcement of an access control decision may take place at many places in a network’s architecture. Access controls may be enforced at network boundaries (e.g., firewalls, routers, and dial-in communications servers) at application servers, or anyplace in the protocol stack or operating system of individual workstations. An important implementation option is to include access control mechanisms at many layers throughout a network architecture.

## 4.3.2 Confidentiality

The confidentiality security service is defined as preventing unauthorized disclosure of data (both stored and communicated). This definition is similar to (and actually a subset of) the description of access control in Section 4.3.1. In fact it can be argued that providing access control also provides confidentiality or conversely, that providing confidentiality is a type of access

control. We include in the definition of “information”, data that is not traditional user data—(examples are network management data, routing tables, password files, and IP addresses on data packets). Confidentiality services will prevent disclosure of data in storage, transiting a local network, or flowing over a public Internet. One subset of confidentiality is that of anonymity, a service that prevents disclosure of information that leads to the identification of the end user.

The requirements for the provision of the confidentiality security service depend on a number of variables:

- *The Location(s) of the Data that Needs Protection*—Data can exist on, an individual machine (e.g., on a hard disk in an end system or in a file on a server) on the wires of a local network; in transport via other mechanisms (e.g., floppy disk) or flowing across a totally public medium (e.g., across the Internet or via a satellite).
- *The Type of Data that Needs Protection*—Data elements may be local files (e.g., passwords, or secret keys) data carried in a network protocol, or the exchanges of a network protocol, (e.g., a protocol data unit).
- *The Different Amounts or Parts of User Data that Need Protection*—It may be necessary to protect an entire data element or only parts of a data element or protocol data unit or the existence of an entire set of protocol exchanges.
- *The Value of the Data that Needs Protection*—The sensitivity and perishability of the data being protected influence the provision of security services, particularly the strength of mechanisms implemented. The value of the data both to the owner and to potential threats must be considered.

The elements of confidentiality are as follows.

- *Data Protection*—This is prevention of disclosure of the contents of data even if it is accessible (e.g., flowing over a network). This element invokes mechanisms that act directly on the data (or act in response to characteristics of the data) rather than acting in response to an entity’s attempt to access data.
- *Data Separation*—Data separation traditionally refers to the concept for providing for separate paths (red/black or physical) or process separation (COMPUSEC techniques, etc).
- *Traffic Flow Protection*—Data characteristics include frequency, quantity, destination of traffic flow, etc. Traffic flow protection includes not only characteristics but inference information such as command structure, and even the instance of communication (e.g., a network communication).

### 4.3.2.1 Data Protection

In cases in which communicated data will be visible to possible adversaries (i.e., via passive monitoring attacks), the most common method for providing confidentiality via data protection is to encrypt the appropriate data. Encryption is also used to protect stored data that might be accessed by an adversary (e.g., via the network-based attacks described in Chapter 3 (Information Systems Security Methodology)).

Encryption is defined as the transformation of data into a form that is unreadable by anyone who does not possess the appropriate secret key. There are many examples of using encryption to provide confidentiality. A small subset includes:

- Security-enabled applications (file encryptors),
- Secure peripherals (media encryptors),
- Operating systems (encrypt local passwords),
- Secure application protocols (secure File Transfer Protocol (FTP)),
- Security protocols (authentication and key management protocols),
- Secure upper layer network protocols (socket layer, IP layer), and
- Secure lower layer network protocols (link encryptors).

There are two types of cryptographic mechanisms that can be used to provide an encryption capability: symmetric cryptography wherein entities which share a common secret key, and public key cryptography (also known as asymmetric cryptography) in which each communicating entity has a unique key pair (a public key and a private key).

Implementation variables in providing encryption for protection of communications data include where in the protocol stack encryption takes place. Encryption at different layers provides different protections to the underlying data or protocol elements.

The strength of the confidentiality service may depend on a number of variables associated with the encryption function.

- The security protocol or application used to invoke the encryption function.
- The trust in the platform executing the protocol or application.
- The cryptographic algorithm.
- The length of the key(s) used for encryption/decryption.
- The protocol used to manage/generate those keys.
- The storage of secret keys (key management keys and encryption keys).

### 4.3.2.2 Data Separation

Data separation takes a different approach to preventing disclosure. Mechanisms that provide data separation prevent the adversary from ever getting at the data in the first place. This is achieved using the normal access control mechanisms described in Section 4.4 (Important Security Technologies), as well as by the additional techniques described below. An example of a commonly used data separation technique is to not allow data labeled as Secret to flow onto an unclassified network.

Data separation mechanisms provide confidentiality by preventing data from reaching a location or destination where it could be disclosed to unauthorized entities. Mechanisms can be employed to prevent data from flowing into undesired areas (routing control). Other mechanisms may be employed to physically segregate those areas. Examples of routing control include a router that directs IP packets based on security labels, thereby preventing Secret packets from reaching unclassified networks, and a firewall that scans e-mail messages for

“dirty words” and prevents messages containing them from being released outside a local network. Examples of physically segregated data are isolated system high networks and physically protected wires.

Data separation mechanisms can be used to counter passive monitoring attacks, as well as insider attacks that inappropriately attempt to release information from a controlled area. The primary variable in the level of assurance provided by a data separation mechanism is the level of trust associated with the process or machine implementing the mechanism.

### 4.3.2.3 Traffic Flow Protection

Data padding can be employed to provide traffic flow protection. Addition of superfluous (usually random) data to data carried in a communications protocol can hide the characteristics (e.g. data rate, data frequency, etc.) of that underlying data. When combined with encryption, this mechanism also hides the content of the underlying data.

Address hiding may also be employed to provide traffic flow protection. Address hiding includes network address translation in which the IP addresses of machines in a local network are replaced by the address of a protecting firewall. Network layer addresses may be hidden by encrypted tunnels, which also provide data confidentiality.

### 4.3.2.4 Other Mechanisms

Other mechanisms for providing confidentiality include spread-spectrum and frequency hopping techniques.

## 4.3.3 Integrity

The integrity security service includes one or more of the following: prevention of unauthorized modification of data (both stored and communicated), detection and notification of unauthorized modification of data, and the journaling of all changes to the data. Modifications to both stored and communicated data may include changes, insertions, deletions, or duplications. Additional potential modifications that may result when data is exposed to communications channels include sequence changes and replay.

The requirements for the provision of the integrity security services are similar to those of confidentiality. These include the location, type, and amount or parts of the data that needs protection.

When integrity is discussed with respect to network security, it is important to consider where in the protocol stack the integrity service is provided. Different implementation (layering) options will provide integrity to different protocol layer data as well as the data being communicated. Sophisticated integrity schemes are likely to require service from the application utilizing the data.

Note that integrity protection is of no value unless it is combined with a mechanism that provides authentication of the source. Without source authentication, anyone could have tampered with the original data and then just reapplied an integrity mechanism.

Data integrity can be divided into two types, based on the type of data to be protected. Integrity can be applied to a single data unit (protocol data unit, database element, file, etc.) or to a stream of data units (e.g., all protocol data units exchanged in a connection).

### 4.3.3.1 Single Unit of Data

Assuring the integrity of a single data unit requires that the originating (sending) entity calculate an additional data item that is a function of (and bound to) the original data unit. This additional item is then carried along with the data unit. The entity that desires to verify the integrity of this data unit must recalculate the corresponding quantity and compare it with the transferred value. A failure of the two to match indicates that the data unit has been modified in transit.

Methods for calculating this piece of data which is a function of the original data unit (the “check value”), vary in the processing required and services provided. Checksums, Cyclic Redundancy Check (CRC) values, and hashes (also known as a message digest) all meet the requirement that they depend on the entire content of the original data unit. A weakness of this method is that, if an adversary modifies the original data, these functions are easily reproducible and allow the adversary to generate a proper value (for the modified data) thereby defeating the integrity service. An additional mechanism can be applied to prevent access to the check value (e.g., encryption or digital signatures) to overcome this problem.

Another method of preventing (successful) modification of the check value is to include a secret value along with the original data unit. This property is exhibited by Message Authentication Codes (also known as Message Integrity Check and keyed hashes).

The icheck value alone will not protect against an attack that replays a single data unit. A time stamp may be included along with the original data unit to provide limited protection against replay.

### 4.3.3.2 Sequence of Data Units

To protect the integrity of a sequence of data units (i.e., protect against reordering, losing, replaying and inserting, or modifying data), some type of ordering information must be provided within the communications protocol. Examples of ordering information are sequence numbers or time stamps. Integrity of sequences can also be provided by encrypting the sequence of data units using a cryptographic algorithm in which encryption of each sequence depends on the encryption of all previous sequences (also referred to as chaining).

## 4.3.4 Availability

Availability is “the property of being accessible and usable upon demand by an authorized entity.” Availability in a networked environment includes not only the user’s ability to access hardware and software resources (such as user agents and servers) but also the user’s ability to obtain a desired Quality of Service (QoS); (e.g., make use of network bandwidth with reasonable throughput). Network traffic must be able to traverse local and wide area networks as required to reach its intended destination.

One of the most effective methods of assuring availability is to provide a secure network environment that exhibits the common security services. Attacks that could prevent a networked system from providing availability may be countered by preventing unauthorized access to resources with access controls and protecting data from disclosure or modification with integrity and confidentiality services. Access control, integrity, and confidentiality become mechanisms to help support the availability security service.

Solutions to problems that affect availability include the following.

- *Protection from Attack*—Some network-based attacks are designed to destroy, degrade, or “crash” network resources. The solution is to harden these resources against such attacks. Means of doing this include closing security holes in operating systems or network configurations, limiting access to resources to authorized entities, and limiting an adversary’s ability to manipulate or view the data flowing through and to those resources (thus preventing insertion of harmful data, such as viruses, or disclosure of sensitive network data, such as routing tables).
- *Protection from Unauthorized Use*—Availability is also limited if a resource is in use, occupied, or overloaded. If unauthorized users are using limited resources (e.g., processing power, network bandwidth, or modem connections), the resources are not available for authorized users. Identifying and authenticating the users of these resources can provide access controls to limit unauthorized use. However, the process of requesting IA to frequently may be used to slow or stop network operations (i.e., non-delivery notice floods).
- *Resistance to Routine Failures*—Normal operational failures and acts of nature also contribute to loss of availability. Solutions include use of equipment designed for high reliability, redundancy in equipment, and network connectivity that provides multiple routes.

Trusted operating system concepts are also used to limit the harmful effects of network attacks. By containing the damage done by malicious code and assuring the proper operation of other security mechanisms, the trusted operating system preserves availability. Another feature exhibited by trusted operating systems is process integrity. This provides assurance that processes executing on an end system provide consistent, repeatable results that are not affected by undesired (unauthorized) influences.

Critical system components must also provide physical security, not only to prevent attacks or misuse of resources, but also to ensure that the platforms and applications are not modified to bypass the invocation of those security services that provide availability.

### 4.3.5 Non-Repudiation

Repudiation is denial by one of the entities involved in a communication that it participated in that communication. The non-repudiation security service provides the ability to prove to a third party that the entity did indeed participate in the communication. When discussed in the context of networking.

- *Non-repudiation with proof of origin* provides the recipient of a data item with proof of the identity of the originator of that data item and the time of origination.

- *Non-repudiation with proof of delivery* provides the originator of a data item with proof that the data item was delivered to the intended recipient (and in some cases, the time of receipt).
- *Auditing services* help provide accountability of the parties involved in exchanges requiring non-repudiation by recording relevant events that can be traceable to persons that can be held responsible for their actions.

The non-repudiation service is primarily provided by application layer protocols. Users are most often concerned with providing non-repudiation for application data (such as an e-mail message or a file). Providing non-repudiation at a lower protocol layer will only provide proof that a particular connection was made; it will not bind the data that flowed over that connection to a particular entity.

Non-repudiation is provided by the authenticating characteristics of digital signatures. A digital signature on a data element (or on the hash of that element) irrevocably ties that data element to the identity contained in the public key certificate associated with the private key that generated the signature. Of course, data integrity must be provided to that data element to assure that the element was not changed after the application of the signature.

Since non-repudiation depends on an identity contained in a public key certificate (and certificates become invalid), it is important to be able to establish, to a third party, the validity of the certificate. It must be possible to prove the validity of that certificate at the time of the original communication as well as any time in the future. This can be accomplished with a combination of trusted time stamps, third party notaries, or archived certificate revocation lists.

Time-stamping achieves the goal of establishing the time at which a communication or transaction occurred. For the highest levels of assurance, time-stamps are applied by a trusted time-stamping service that digitally signs the data item (or a hash of the data item) along with the time stamp before delivery to the intended recipient.

## 4.4 Important Security Technologies

An overview of technical security countermeasures would not be complete without at least a high-level description of the widely used technologies underlying those countermeasures. This section highlights selected technologies as an introduction to the detailed technology assessments included in Chapters 5 through 9 (Defend the Network and Infrastructure, Defend the Enclave Boundary/External Connections, Defend the Computing Environment, Supporting Infrastructures, and Information Assurance for the Tactical Environment). For convenience, these technologies are listed alphabetically.

- **Application Layer Guard**—The need for a separate mechanism to perform a gatekeeper function, checking the invocation of security features, gives rise to a need for security at the application layer. This gatekeeper has recently taken the form of an application layer guard that implements firewall mechanisms (performing I&A functions and enforcing security policies such as allowing or disallowing connections based on identification and/or requested protocol processing). Guard functionality includes such features as performing a cryptographic invocation check on information that is allowed outside the protected enclave, and data content filtering to support sensitivity regrade

decisions. The guard functionality, while effective for non-real-time applications (e.g., e-mail) on networks with low sensitivity, has been difficult to scale to highly classified networks and real-time applications.

- **Application Program Interface (API)**—APIs are a means of isolating a computing platform from the details of the implementation of cryptographic functions (both the actual algorithms and the hardware implementations). It provides standard interfaces so that multiple vendors may provide interoperable solutions.
- **Common Data Security Architecture (CDSA)**—The CDSA is a set of layered security services that address communications and data security problems in the emerging Internet and intranet application space. CDSA focuses on security in peer-to-peer distributed systems with homogeneous and heterogeneous platform environments. The architecture also applies to the components of a client/server application. The CDSA addresses security issues and requirements in a broad range of applications by:
  - Providing layered security mechanisms (not policies);
  - Supporting application-specific policies by providing an extensibility mechanism that manages add-in (policy-specific) modules;
  - Supporting distinct user markets and product needs by providing a dynamically extensible security framework that securely adds new categories of security service;
  - Exposing flexible service provider interfaces that can accommodate a broad range of formats and protocols for certificates, cryptographic keys, policies, and documents; and
  - Supporting existing, secure protocols, such as Secure Sockets Layer (SSL), Secure/Multipurpose Internet Mail Extension (S/MIME), and Secure Electronic Transaction (SET).
- **Circuit Proxy**—Circuit gateways are another type of proxy firewall. A circuit-level proxy becomes an intermediate connection point in a session between a client and a server. To reach a distant server, a client initially connects to a TCP port on the circuit proxy machine. The circuit proxy then completes the connection (after making an access control decision) to the target server. Access controls are based on the identity of the initiating machine without interpreting the application protocol or viewing the contents of protocol packets. A circuit-level proxy can be utilized across several application protocols; however, client modifications may be necessary to use the circuit-level protocol.
- **CryptoAPI**—The Microsoft Cryptographic API provides services that enable application developers to add cryptography to their Win32 applications. Applications can use the functions in CryptoAPI without knowing anything about the underlying implementation, in much the same way that an application can use a graphics library without knowing anything about the particular graphics hardware configuration.
- **Cryptographic Service Providers (CSP)** —Both CDSA and CryptoAPI make use of the concept of CSPs which are independent modules that perform the real cryptographic work. Ideally, CSPs are written to be completely independent of any particular application, so that a given application will run with a variety of CSPs. In reality, however, some applications may have very specific needs that require a custom CSP.



A CSP may implement one or more of the following cryptographic functions: bulk encryption algorithm, digital signature algorithm, cryptographic hash algorithm, unique identification number, random number generator, secure key storage, and custom facilities unique to the CSP.

A CSP may be implemented in software, hardware, or both. A CSP or an independent module can also deliver key management services, such as key escrow or key recovery. CSPs should not reveal key material unless it has been wrapped. Also, the key-generation function of a CSP should be made as tamper-resistant as possible.

- **File Encryptors**—These provide confidentiality and integrity for individual files, provide a means of authenticating of a file's source, and allow the exchange of encrypted files between computers. File encryptors typically implement a graphical user interface (GUI) that allows users to choose files to be encrypted or decrypted. This protects individual files, but it does not protect all of the files on the drive.

Many applications generate temporary files that may contain user data. These files are normally erased when the application is closed; but when the application does not close in an orderly fashion, these temporary files may remain. Some operating systems do not actually erase data when files are deleted. Instead, they alter the name of the file in the file allocation table. The user's data remains on the hard drive until the space is reallocated to another file and overwritten. Thus, unencrypted and potentially classified user data can remain on the hard drive after system shutdown, either by failure to erase temporary files or by design of the operating system's erasing function.

- **Hardware Tokens**—A number of hardware token approaches are available ranging from an external memory device to one with significant levels of processing. One hardware token that is prominent in the Department of Defense (DoD) community is the FORTEZZA® Crypto Card. The FORTEZZA® card provides the cryptographic algorithms required to provide security services to a FORTEZZA® based system. It stores the private key information for each user personality, the certificates of its issuers, and the public keys needed for cryptography. It performs the digital signature and hash algorithms, public or private key exchange functions, encryption, and decryption. The interface to the card depends on the hardware platform and its configuration, and the operating system.
- **Intrusion and Penetration Detection**—Intrusion detection and response systems can protect either a network or individual client platforms. Effective intrusion systems detect both insider and outsider threats. In general, intrusion systems are intended to protect against and respond to situations in which the available countermeasures have been penetrated, either through allowed usage or the exploitation of vulnerabilities that are unknown or have not been patched. The objective of these systems is to detect malicious and unintended data and actions (e.g., altered data, evil executables, requests that permit unintended resource access, and unintended use of intended services). Once the intrusion is detected, an appropriate response is initiated (e.g., disconnect attacker, notify operator; respond automatically to halt or lessen the attack; trace attack to proper source; and counter the attack, if appropriate). Intrusion detection mechanisms operating at the transport layer can view the contents of transport packets (e.g., TCP packets) and are able to detect more sophisticated attacks than mechanisms that operate at the network layer. Intrusion detection mechanisms operating at the network layer can view

the contents of network packets (e.g., IP packets) and are thus only able to detect attacks that are manifested at the network layer (e.g. port scans).

- **Internet Security (IPSec)** —IPSec is the security framework standardized by the IETF as the primary network layer protection mechanism. IPSec consists of two parts; an Authentication Header (AH), whose purpose is to bind the data content of IP frames to the identity of the originator and an Encapsulating Security Payload (ESP) for privacy. The authentication header is intended for use when integrity of information is required but privacy is not. ESP is intended for use where data confidentiality is required. ESP defines two methods (or modes) of encapsulating information. Tunnel mode, when used at an enclave boundary, aggregates traffic flow from site to site and thereby hides end system identification. Transport mode leaves end system identification in the clear and is most advantageous when implemented at the end system.
- **Internet Key Exchange (IKE) Protocol**—IKE was developed by the IETF as a standard for security attribute negotiation in an IP network. It provides a framework for creating security associations between endpoints on an IP network, as well as the methodology to complete the key exchange. IKE is based upon the Internet Security Association Key Management Protocol (ISAKMP) with OAKLEY extensions. The structure of ISAKMP is sufficiently flexible and extensible to allow inclusion of future security mechanisms and their associated algorithms and can be tailored to other networking technologies.
- **Media Encryptors**—Media encryptors protect the confidentiality and integrity of the contents of data storage media. They can also perform a role in maintaining the integrity of the workstation by verifying the Basic Input/Out System (BIOS) and ensuring that configuration and program files are not modified. Media encryptors need to leave some system files unencrypted so that the computer can boot from the hard drive. Most of these files can have their integrity protected by a cryptographic checksum; this will not prevent a tamper attack, but it will alert the user that the data has been altered. However, some system files contain data that changes when the computer is booted; these files cannot be protected. With the exception of some system files, media encryptors encrypt the entire contents of the drive.
- **Packet Filter**—Packet filtering firewalls (also referred to as screening routers) commonly operate at the network layer (Open Systems Interconnection (OSI) Layer 3). These firewalls examine the IP and protocol headers against a set of predefined rules. They can typically filter packets based on host and destination IP address, port number, and the interface. This type of firewall is generally inexpensive, fast, and transparent to the user. However, screening routers generally do not have a very robust auditing capability, nor do they allow the use of strong authentication on incoming connections. The combination of a packet filtering system and another product (authentication server) may provide strong authentication capability.
- **PKI Certificate Management Protocol**—For managing public key material, the Internet community has developed the Internet X.509 Public Key Infrastructure Certificate Management Protocols (CMP). Management protocols are required to support on-line interactions between PKI components. For example, a management protocol might be used between a CA and a client system with which a key pair is associated, or between two CAs that cross-certify each other. At a high level, the set of operations for which management messages are defined can be grouped as follows.

- CA Establishment. When establishing a new CA, certain steps are required (e.g., production of initial CRL, export of CA public key).
  - End Entity Initialization. This includes importing a root CA public key and requesting information about the options supported by a PKI management entity.
  - Certification. Various operations result in the creation of new certificates:
    - Initial registration/certification,
    - Key pair update,
    - Certificate update,
    - CA key pair update,
    - Cross-certification, and
    - Cross-certificate update.
  - Certificate/CRL Discovery Operations. Some PKI management operations result in the publication of certificates or CRLs:
    - Certificate publication, and
    - CRL publication.
  - Recovery Operations. Some PKI management operations are used when an end entity has “lost” its key material.
  - Revocation Operations. Some PKI operations result in the creation of new CRL entries and/or new CRLs.
- SSL—SSL exists just above the transport layer and provides security independent of application protocol although its initial implementation was meant to secure the Hypertext Transfer Protocol (HTTP). This effort has migrated to the IETF as the Transport Layer Security (TLS) protocol. The TLS protocol provides data encryption, server authentication, message integrity, and optional client authentication for a TCP/IP connection. It negotiates the invocation of cryptographic algorithms (from a fixed set) and protects all application layer data.
  - S/MIME—S/MIME is a specification for adding security for e-mail in Multipurpose Internet Mail Extensions format, supporting binary attachments as well as text. It offers authentication and confidentiality. S/MIME uses a hybrid approach to providing security, referred to as a digital envelope. The bulk message is encrypted with a symmetric cipher, a public key algorithm is used for key exchanges and for digital signatures, and X.509 certificates support authentication. S/MIME supports anonymity to the extent that it applies the digital signature first, and then encloses the signature and the original message in an encrypted digital envelope, so that no signature information is exposed to a potential adversary.

The S/MIME specification is currently an Internet draft that recommends three symmetric encryption algorithms: Data Encryption Standard (DES), Triple-DES, and RC2 (a symmetric block cipher with a 40-bit key to meet the U.S. Government export requirements). It also builds on the Public Key Cryptography Standards (PKCS), specifically PKCS #7, providing a flexible and extensible message format for representing the results of cryptographic operations, and PKCS #10, a message syntax for certification requests. The S/MIME specification has been submitted to the IETF in an effort to make it an industry-accepted standard.

- **SOCKS**—This protocol supports application-layer firewall traversal. The SOCKS protocol supports both reliable TCP and User Datagram Protocol (UDP) transport services by creating a shim-layer between the application and transport layers. The SOCKS protocol includes a negotiation step whereby the server can dictate which authentication mechanism it supports. Compliant implementations must support Generic Security Services (GSS)-API and username/password authentication modes.
- **Stateful Packet Filter**—Stateful packet filters look at the same headers as do packet filters, but also examine the content of the packet. In addition, this technology is capable of dynamically maintaining information about past packets or state information. Security decisions can then be based on this state information. Because they have the ability to retain state information, stateful packet filters permit UDP-based services (not commonly supported by firewalls) to pass through the firewall. Thus they are advertised as offering greater flexibility and scalability. Stateful packet filtering technology also allows for logging and auditing and can provide strong authentication for certain services.
- **Trusted Computing Base (TCB)** —A Trusted Computer System is a system that employs sufficient hardware and software assurance measures to allow its use for simultaneous processing of a range of sensitive or classified information. Such a system is often achieved by employing a TCB. A TCB is the totality of protection mechanisms within a computer system, including hardware, firmware, and software, the combination of which is responsible for enforcing a security policy. A TCB consists of one or more components that together enforce a unified security policy over a product or system. The ability of a TCB to correctly enforce a unified security policy depends solely on the mechanisms within the TCB and on system administration personnel's correct input of parameters (e.g., a user's clearance level) related to the security policy.
- **Virus Detectors**—Virus detectors can be employed to protect a network or an individual client. A virus can be considered a special form of intrusion involving the classical Trojan horse attack with the ability to reproduce and spread. The virus is normally considered to be limited to the authorizations of the user who is executing the code, but viruses may also exploit flaws in the network that allow the virus to cause a serious privilege state harm.

## 4.5 Robustness Strategy

### Purpose

The robustness strategy describes a process that, when completed in a later release of the IATF, will provide guidance in assessing the *degree of robustness*. Robustness is defined as the level of security mechanism strength and assurances recommended (considered “good enough”) in an Information Security (INFOSEC) solution. At the current stage of development, the strategy deals primarily with the levels within individual security services and mechanisms, based on information on a given value, in a particular (static) threat environment. As discussed below, this is not a complete answer. The process is not intended to provide an endorsement or credential for specific products, nor is it intended to serve as a “cookbook” answer for the robust-

ness of solutions; rather, it offers security engineering guidance to developers, integrators, and risk managers as input to risk management. Users of the IATF can employ the robustness strategy for:

- Providing guidance to help developers and integrators assess what strength of mechanisms what levels of assurance (in development methodology, evaluation, and testing); and (3) what criteria are recommended for a particular configuration meant to protect information of a particular value, with a specific intelligence life, in a specific, static threat environment;
- Defining product requirements for different customer scenarios (value of information, threat, configuration, etc.) for example as described in the IATF;
- Providing feedback to security requirements developers, decision-makers, customer representatives, customers, etc.;
- Constituting developmental requirements when a security solution does not exist;
- Working with academia to foster research in the network security arena, and to educate future engineers, architects, and users in network security technology; and
- Performing subsequent risk assessments made necessary by reconfiguration of the system/network under review or by a change in threat or value of information.

As technology in general and INFOSEC threats in particular evolve, countermeasures will need to evolve, and with them the corresponding application guidance. This paper is a strategy for the development of a general security mechanism/countermeasure valuation scheme. Rather than directly defining the security requirements, that need to be met, it characterizes the relative strength of mechanisms, that provide security services and provides guidance in selecting these mechanisms.

Trained ISSEs [11] support customer organizations in defining and applying security solutions to address their Information Assurance (IA) needs. Working with a customer from initial contact through solution acceptance, an ISSE helps ensure that the customer's security needs are appropriately identified and that acceptable solutions are developed. Within the context of the IATF robustness strategy, an ISSE helps the customer assess the value of his or her information and assets and the security threat within the operational environment, identifies the security services necessary to provide appropriate protection, and provides guidance on the characteristics of the specific security mechanisms that provide those services.

Multiple applications of the same system/environment but with differently trained ISSEs may result in different guidance, and though all such outcomes would be consistent with the recommended use of the strategy. There is no concept of official compliance with the robustness strategy in terms of approving a solution. It is a strategy, that is an aid to "getting you there" as opposed to a prescriptive solution (where nominal compliance assures acceptability).

## Robustness Strategy Section Overview

The Overview of the General Process section describes the general process, including assumptions and output. Section 4.5.2, *Determining the Degree of Robustness* section presents an approach for determining recommended robustness (strength of mechanism and assurance) levels based on the value of information to be protected and the threat environment. Section 4.5.3,

*Strength of Mechanism*, breaks down security services into supporting mechanisms and identifies corresponding strength levels. The *Level of Assurance* section (Section 4.5.4) discusses related aspects of obtaining assurance. A *Worked Example* (Section 4.5.5) demonstrates how the process would be applied in developing specific guidance. This is followed by a discussion of *Robustness Strategy Evolution* (Section 4.5.6) which provides recommendations for those who would carry on the work outlined in this paper. Then lastly, Section 4.5.7, *Real World Applications*, which demonstrates real world application of the Robustness Strategy.

## 4.5.1 Overview of the General Process

The robustness strategy is intended for application in the context of the development of a security solution and is meant to be consistent with IATF Chapter 3 (Information System Security Engineering), which describes the overall process. An integral part of that process is determining the recommended strength and degree of assurance of proposed security services and mechanisms that become part of the solution set. The strength and assurance features serve as a basis for the selection of the mechanisms and as a means of evaluating the products that implement those mechanisms. This section provides guidance on determining the recommended strength and assurance.

The process should be applied to all components of a solution, both products and systems, to determine the robustness of configured systems as well as of their component parts. It applies to Commercial Off-The Shelf (COTS), Government Off-The-Shelf (GOTS), and hybrid solutions. As indicated above, the process is to be used by security requirements developers, decision-makers, ISSEs, customers, and others involved in the solution life cycle. Clearly, if a solution component is modified, or threat levels or the value of information changes, there must be a reassessment of risk with respect to the new configuration.

Various risk factors, such as degree of damage suffered if the security policy is violated, threat environment, and so on, will be used to guide determination of an appropriate strength, and associated level of assurance for each mechanism. Specifically, the value of information to be protected and the perceived threat environment are used to obtain guidance on the recommended Strength of Mechanism Level (SML) and Evaluation Assurance Level (EAL).

## 4.5.2 Determining the Degree of Robustness

We define the degree of robustness as the level of strength and assurance recommended for potential security mechanism(s). To determine this level for a given security service in a particular application, the customer and the ISSE should consider the value of the information to be protected (in relation to the operational mission) as well as the perceived threat environment. Guidelines for determining these values are provided below. Once a determination has been made regarding the information value and threat environment, the ISSE uses the Robustness Table 4-7 to determine required assurance and strength of mechanism levels.

Note that the robustness strategy focuses specifically on individual security services and mechanisms. When the robustness of an overall network solution is considered, the individual solutions at each layer within the network must also be considered. IA mechanisms can be applied at the host, sub-net, boundary, and backbone levels. Robustness should take into account

the implications of composing layered protection mechanisms and also incorporates an overall assessment of vulnerabilities and residual risks for each layer.

Many customers, in support of their mission, have a need to protect information (or an information system) whose compromise could adversely affect the security, safety, financial posture, or infrastructure of the organization. Five levels of information value have been defined:

- *V1*: Violation of the information protection policy would have negligible adverse effects or consequences.
- *V2*: Violation of the information protection policy would adversely affect and/or cause minimal damage to the security, safety, financial posture, or infrastructure of the organization.
- *V3*: Violation of the information protection policy would cause some damage to the security, safety, financial posture, or infrastructure of the organization.
- *V4*: Violation of the information protection policy would cause serious damage to the security, safety, financial posture, or infrastructure of the organization.
- *V5*: Violation of the information protection policy would cause exceptionally grave damage to the security, safety, financial posture, or infrastructure of the organization.

Similarly, the customer must work with an ISSE to define the threat environment in which the mission will be accomplished. Things to consider when determining the threat to a particular solution include level of access, risk tolerance, expertise, and available resources obtainable by the adversary. These threats should be considered in the context of the system security policy.

The following threat levels were derived from various relevant works (e.g., Security Management Infrastructure (SMI) Task 1 Team, Threat and Vulnerability Model for Information Security, 1997 [12]), and discussions with subject matter experts throughout the Information Systems Security Organization (ISSO). Seven levels of threat have been defined.

- *T1*: Inadvertent or accidental events (e.g., tripping over the power cord).
- *T2*: Passive, casual adversary with minimal resources who is willing to take little risk (e.g., listening).
- *T3*: Adversary with minimal resources who is willing to take significant risk (e.g., unsophisticated hackers).
- *T4*: Sophisticated adversary with moderate resources who is willing to take little risk, e.g., organized crime, sophisticated hackers, international corporations.
- *T5*: Sophisticated adversary with moderate resources who is willing to take significant risk (e.g., international terrorists).
- *T6*: Extremely sophisticated adversary with abundant resources who is willing to take little risk (e.g., well-funded national laboratory, nation-state, international corporation).
- *T7*: Extremely sophisticated adversary with abundant resources who is willing to take extreme risk, (e.g., nation-states in time of crisis).

After a determination is made regarding the value of the information to be protected and the threat environment, the ISSE can provide guidance on how strong the security mechanism should be and what assurance activities that should be performed. Table 4-7 indicates the minimal recommended SML and EAL[6] or providing protection of information or information

systems of a given value (V1toV5) against a given threat level (T1toT7). Section 4.5.3 (Strength of Mechanism) defines the SMLs and Section 4.5.4 (Level of Assurance) defines the EALs.

**Table 4-7 Degree of Robustness**

Information Value	Threat Levels						
	T1	T2	T3	T4	T5	T6	T7
<b>V1</b>	SML1 EAL1	SML1 EAL1	SML1 EAL1	SML1 EAL2	SML1 EAL2	SML1 EAL2	SML1 EAL2
<b>V2</b>	SML1 EAL1	SML1 EAL1	SML1 EAL1	SML2 EAL2	SML2 EAL2	SML2 EAL3	SML2 EAL3
<b>V3</b>	SML1 EAL1	SML1 EAL2	SML1 EAL2	SML2 EAL3	SML2 EAL3	SML2 EAL4	SML2 EAL4
<b>V4</b>	SML2 EAL1	SML2 EAL2	SML2 EAL3	SML3 EAL4	SML3 EAL5	SML3 EAL5	SML3 EAL6
<b>V5</b>	SML2 EAL2	SML2 EAL3	SML3 EAL4	SML3 EAL5	SML3 EAL6	SML3 EAL6	SML3 EAL7

Using an applicable Capability Maturity Model (CMM), Capability Level 2 or equivalent is recommended for EALs 1to3 and a Capability Level 3 or equivalent for EALs 4-7.

The Systems Security Engineering Capability Maturity Model (SSE-CMM) is one example of an applicable CMM. A capability maturity model describes the stages through which processes advance as they are defined, implemented, and improved.<sup>1</sup> The SSE-CMM is designed to support a host of improvement activities, including self administered appraisals or internal appraisals augmented by experts (e.g., ISSEs) from inside or outside of the organization.<sup>2</sup>

The ISSE, working with the customer, would apply the SSE-CMM (or another applicable CMM) as a baseline capability. The assessment of compliance is still left to the discretion of the customer. Reasonable justification is still necessary and it should be denoted that acquisition personnel need to be knowledgeable about the CMM used.

### 4.5.3 Strength of Mechanism

SML is presented by a series of tables focusing on specific security services. The strategy is still being formulated, and the tables are not considered complete or adequately refined. There are a number of additional security mechanisms that are not detailed in the tables but that may be appropriate for providing some security services. Further, the strategy is not intended, by itself to provide adequate information for the selection of the desired (or sufficient) mechanisms for a particular situation. As indicated earlier, an effective security solution will only result from

<sup>1</sup> System Security Engineering Capability Maturity Model Description document

<sup>2</sup> System Security Engineering Capability Maturity Model Summary



the proper application of ISSE skills to specific operational and threat situations. The strategy does offer a methodology for structuring a more detailed analysis. The security services itemized in these tables have several related supporting security services that may result in recommendations for inclusion of additional security mechanisms and techniques.

For each service, recommended guidance on each SML levels is given for a variety of mechanisms that provide the overall service. In some cases, a group of mechanisms will be required to provide the necessary protection. It should also be noted that an ISSE, in conjunction with a customer, could decide to use a stronger or weaker mechanism than is recommended, depending on the environment. It is the intent of the strategy to ensure that mechanisms across services at the same strength level provide comparable protection, in that they counter equivalent threats. The selection of mechanism(s) from service tables is an independent event, in the sense that one mechanism does not necessarily require others. Higher strength mechanisms do not necessarily contain features of lower strength mechanisms (i.e., security functions do not necessarily accumulate at higher strength levels). Table entries are preliminary estimates based on consultation with subject matter experts and are likely to be revised based on technology evolution, threat assessment, and costing development.

The strength referred to below is a *relative* measure of the effort (cost) required to defeat the mechanism and is not necessarily related to the cost of implementing such countermeasures. All things being equal, (especially cost), the highest strength mechanism should always be chosen. Three SMLs are defined:

- *SML1* is defined as basic strength or good commercial practice. It is resistant to the unsophisticated threat (roughly comparable to the T1toT3 threat levels) and is used to protect low-value data. Examples of countered threats might be door rattlers, ankle biters, inadvertent errors.
- *SML2* is defined as medium strength. It is resistant to the sophisticated threat (roughly comparable to the T4toT5 threat levels) and is used to protect medium-value data. It would typically counter a threat from an organized effort (e.g. an organized group of hackers).
- *SML3* is defined as high strength or high grade. It is resistant to the national laboratory or nation-state threat (roughly comparable to the T6toT7 threat levels) and is used to protect high-value data. An example is an extremely sophisticated, well-funded technical laboratory or a nation-state adversary.

Based on these definitions, the customer and ISSE will apply their knowledge of the specific operational and threat situation to determine what strength of mechanism is recommended for each of the mechanisms listed in the following sections.

### 4.5.3.1 Mechanisms Supporting Security Management

Recommended mechanisms for establishing needed security management are depicted in Table 4-8. The degree of awareness and control with respect to the following will identify the SML target.

- *Compromise recovery*, in addition to achieving a secure initial state, secure systems must have a well-defined status after failure, either to a secure failure state or via a recovery procedure to a known secure state.
- *Poor system administration* is a leading cause of security weaknesses and vulnerabilities. It is the first line of defense in enforcing the security policy. (See IATF Chapter 3 Information Systems Security Engineering) for more information on system security administration.
- *Training* is what operators and users need to obtain to learn about security features and system operation. Knowledgeable users are more likely to exercise due care in protecting information assets (increased risk of insider attack is dealt with via by means of personnel security).
- *The Operational Security (OPSEC)* process is a coordinated, multidisciplinary five-step activity involving identification of critical information, threat identification and analysis, vulnerability identification and analysis, risk assessment, and adoption of countermeasures. Each use of the process is tailored to a specific activity of concern, which is examined for potential disclosure to specific adversaries, upon which to base directly pertinent countermeasures. Consult with the Interagency Operation Support Staff for consideration of individual cases.
- *Trusted distribution* is a calculated/controlled method of distributing security-critical hardware, software, and firmware components. It protects of the system from modification during distribution and detects any changes.
- *Secure operations* is the level of standard operating procedures needed to provide security given the classification, sensitivity, and criticality of the data and resources being handled or managed. This includes security doctrine.
- *Mechanism management*, certain security mechanisms (e.g., cryptographic algorithms) have ancillary support needs (e.g., key management).

**Table 4-8 Security Management Mechanisms**

	<b>Compromise Recovery</b>	<b>System Administration</b>	<b>Training</b>	<b>OPSEC</b>	<b>Trusted Distribution</b>	<b>Secure Operations</b>	<b>Mechanism Management</b>
<b>SML1</b>	Informal plan	See Ch. 4 counter-measures	Training available at user discretion	Implement OPSEC at user's discretion	Direct vendor purchase	Informal plan of operation	Procedural, user's discretion
<b>SML2</b>	Detailed plan that is reviewed and approved	See Ch. 4 counter-measures	Formal training plan	OPSEC training required; implement at user's discretion	Certificate of authenticity, virus scan, validation	Formal plan of operation	Procedural, reminders, user's discretion
<b>SML3</b>	Detailed plan that is reviewed and approved	See Ch. 4 counter-measures	Knowledge/skill certification required	OPSEC training required, implementation required	Protective packaging, checksums, validation suite	Detailed, formal plan of operation	Automated support

### 4.5.3.2 Mechanisms Supporting Confidentiality

Confidentiality is the protection of information against disclosure to unauthorized entities or processes. Possible security mechanisms for this security service are depicted in Table 4-9. These mechanisms can be obtained individually or in combination.

- *If cryptographic algorithm* is chosen, some of the factors that must be considered are the management of keying material and the effective length of the key, which includes the strength of the underlying cryptographic algorithm. Effective key length is defined as the nominal key length, reduced by the effect of any known attacks against the cryptographic algorithm (assuming correct implementation). The supporting KMI [9] categories are defined in Chapter 8 (Supporting Infrastructures).
- *Physical security* includes tangible security mechanisms such as guards, locks, and fences. The idea is to build a physically secure enclave, providing guards and high walls.
- *Technical security* is a protection mechanism for hardware. Tampering is the unauthorized modification that alters the proper functioning of an information security device or system in a manner that degrades the security or functionality it provides. Anti-Tamper mechanisms detect such alterations. TEMPEST is the investigation, study, and control of compromising emanations from telecommunications and Automated Information System (AIS) equipment.
- *Anonymity* is the desire for a user to remain unknown during a virtual transaction. Some applications requiring anonymity might be Internet voting and Internet cash. This area is relatively immature and is currently addressed by the Transmission Security (TRANSEC)[10] and cover & deception disciplines. TRANSEC mechanisms provide various degrees of covertness to prevent detection, identification and exploitation. Cover and deception can be provided through such mechanisms as anonymous remailers, “onion routing”, or “web anonymizers.” Cover and deception currently has no differentiated levels.

**Table 4-9 Confidentiality Mechanisms**

	Cryptographic Algorithm		Physical Security	Technical Security		Anonymity	
	Effective Key Length	Key Management		Anti tamper	TEMPEST	TRANSEC	Cover & Deception
<b>SML1</b>	40+ bits symmetric key length, 80+ exponent 512+ modulus public key length	SMI Cat X, 80+ exponent 512+ modulus public key length, 80+ hash key length	Comparable to [7]	[6] level 1 or 2	comply with applicable EMI/EMC FCC standards or portions of [8]	Low power unit	TBD
<b>SML2</b>	80+ bits symmetric key length, 160+ exponent 1024+ modulus public key length	SMI Cat Y, 160+ exponent 1024+ modulus public key length, 160+ hash key length	Comparable to [7]	[6] level 3 or 4	[8]	Commercial spread spectrum signal techniques	TBD
<b>SML3</b>	Because of the complicated nature of this level, please consult with a qualified ISSE. <sup>3</sup>	SMI Cat Z, also consult with a qualified ISSE. <sup>3</sup>	Comparable to [7]	[6] level 4 or better	[8]	cryptographic spread spectrum signal techniques	TBD

### 4.5.3.3 Mechanisms Supporting Integrity

In Table 4-10 there are four mechanisms that will help in ensuring integrity, either singly or in combination with others. When taken in the context used here, integrity, as a security service, means the protection of information against undetected, unauthorized modification, or undetected destruction of information.

- A cryptographic algorithm in an error extension mode will emphasize the error and should be used in conjunction with a detection mechanism (e.g., parity or human review).
- Physical security is described in Table 4-9.
- Signature/Checksum provides data integrity by digitally signing data. Typically, the data requiring protection is used to calculate a smaller value, such as a parity, checksum, or hash. This value can then be digitally signed.

<sup>3</sup> DoD users should consult with a National Security Agency ISSE. Other government users are directed to contact an ISSE at the National Institute of Standards and Technology for guidance in this area. Non-government users should consult with a qualified ISSE, or equivalent representative within their organization.

- Redundancy is the availability of multiple methods to obtain the same information.

**Table 4-10 Integrity Mechanisms**

	Cryptographic Algorithm		Physical Security	Signature Checksum	Redundancy
	Effective Key Length	Key Management			
<b>SML1</b>	40+ bits symmetric key length, 80+ exponent 512+ modulus public key length	SMI Cat., 80+ exponent 512+ modulus public key length, 80+ hash key length	comparable to [7]	Parity, or commercial checksum, hash and signature with SML1 algorithm	Not applicable
<b>SML2</b>	80+ bits symmetric key length, 160+ exponent 1024+ modulus public key length	SMI Cat, 160+ exponent 1024+ modulus public key length, 160+ hash key length	comparable to [7]	Cryptographic checksum, hash, and signature with SML2 algorithm	Redundant data path with 100% correct comparison
<b>SML3</b>	Due to the complicated nature of this level, please consult with a qualified ISSE. <sup>4</sup>	SMI Cat, also consult with a qualified ISSE. <sup>4</sup>	comparable to [7]	Cryptographic checksum, hash and signature with SML3 algorithm	Multiple data paths with 100% correct comparison

### 4.5.3.4 Mechanisms Supporting Availability

Availability is also known as service assurance. To ensure availability of data, the system must employ both preventive and recovery mechanisms. This security service is quantified in Table 4-11 and can be obtained through a combination of the services as appropriate for the applications.

- TRANSEC is used to overpower potential jammers. A strong enough signal is provided for this anti-jam capability. TRANSEC can also be used to hide a signal to avoid jamming. (Note that, because of the real-time nature of exploitation, it may not be necessary to use an SML3 algorithm strength to meet the SML3 level for this mechanism).
- Anti-tamper mechanism is described in Table 4-9.
- Physical security is described in Table 4-9.
- Redundancy or redundant paths should be available to allow information flow without violating the site security policy. Such information flow might include bypassing any problem areas, including congested servers, hubs, cryptography and so on.
- Data recovery is the ability to recover data that might otherwise be unavailable due to the loss of key, storage media, etc.

<sup>4</sup> DoD users should consult with a National Security Agency ISSE. Other government users are directed to contact an ISSE at the National Institute of Standards and Technology for guidance in this area. Non-government users should consult with a qualified ISSE or an equivalent representative within their organization.

**Table 4-11 Availability Mechanisms**

	TRANSEC	Anti-Tamper	Physical Security	Redundancy	Data Recovery
<b>SML1</b>	High power	level 1 or 2 [4]	comparable to [7]	Bypass channel available	Informal archival plan, user backs up own key or data
<b>SML2</b>	Commercial spread spectrum signal techniques	level 3 or 4 [4]	comparable to [7]	Backup data path, hot spare	Formal archival plan, central back-ups
<b>SML3</b>	Cryptographic spread spectrum signal techniques	level 4 or better [4]	comparable to [7]	Multiple data paths, multiple hot spares	Formal archival plan, central, offsite back-ups

### 4.5.3.5 Mechanisms Supporting Identification and Authentication (I&A)

I & A is required for effective access control. There usually is a need for a process that enables recognition of an entity within or by an AIS. Along with that, a security measure designed to establish the validity of a transmission, message, or originator or a means of verifying an individual's eligibility to receive specific categories of information is needed. These are the attributes of I&A that are listed in Table 4-12. We categorize these attributes as follows.

- *Identification or System Identification (SID)* in particular is one way in which a system might recognize the “entity” (which may be a person requesting authentication). Biometrics might be used to identify a living person.
- *Human-to-machine authentication* could utilize alphanumeric phrases, like passwords, Personal Identification Numbers (PIN), or challenge/response exchanges that are memorized by a human or used with a token calculator (e.g. challenge/response). Also, physical devices, such as hardware tokens, have this utility (e.g., a credit card-type physical entity).
- *Peer-to-peer authentication* can utilize certificates that identify and authenticate the entities. Along with the certificate is the similar SML cryptographic algorithm that “binds” it to the entity with a digital signature. Authentication is provided by a trusted third party (a separate, but knowledgeable entity). Within this area, one could use a cryptographic algorithm (as discussed under confidentiality above), and personnel security policy, where a security clearance is obtained for a particular person to reduce the risk of an insider attacking the system.

**Table 4-12 Identification and Authentication Mechanisms**

	Identification		Human-to-Machine Authentication		Peer-to-Peer Authentication			
	System IDs	Bio-metrics	Passwords PINS Challenge/Response	Tokens	Certificates	Cryptographic Algorithm		Personnel Security
						Effective Key Length	Key Management	
<b>SML1</b>	Uniqueness	Not applicable	Have one	Badge/key static	bind w/SML1 cryptographic algorithm	40+ bits symmetric key length, 80+ exponent 512+ modulus public key length	SMI Cat. X, 80+ exponent 512+ modulus public key length, 80+ hash key length	Commercial hiring practices
<b>SML2</b>	Uniqueness and minimum character length	Use one Biometric	Minimum effective length - TBD	Memory device, updated periodically	bind w/SML2 cryptographic algorithm	80+ bits symmetric key length, 160+ exponent 1024+ modulus public key length	SMI Cat Y, 160+ exponent 1024+ modulus public key length, 160+ hash key length	Equivalent of Secret clearance
<b>SML3</b>	Uniqueness and minimum character length, minimum distance (e.g., Hamming)	Use one Biometric with a liveness test	Minimum effective length - TBD	CIK, updated every time	bind w/SML3 cryptographic algorithm	Because of the complicated nature of this level, please consult with a qualified ISSE. <sup>5</sup>	SMI Cat Z, also consult with a qualified ISSE. <sup>5</sup>	Equivalent of Top Secret clearance

### 4.5.3.6 Mechanisms Supporting Access Control

Beyond I&A, access control can be thought of as a “super service” encompassing all security services. In the context of network security, access control is concerned with limiting access to networked resources (hardware and software) and data (stored and communicated). The primary goal here is to prevent unauthorized use, and unauthorized disclosure or modification of data by unauthorized entities. A secondary goal is to prevent an availability attack (e.g., denial-of-service attack). Several mechanisms that can be used to help provide the access control service are shown in Table 4-13 and include the following parameters.

- *Anti-tamper* is described under Confidentiality (Table 4-9).

<sup>5</sup> DoD users should consult with a National Security Agency ISSE. Other government users are directed to contact an ISSE at the National Institute of Standards and Technology for guidance in this area. Non-government users should consult with a qualified ISSE, or an equivalent representative within their organization.

- *Mandatory Access Control (MAC)* is where authorized access to data is automatically imposed by the system through the use of labels and binding the labels to the data associated with it. When implementing MAC, there is a concern with the integrity of the label itself and the strength of binding of the label to the data. In other words, if SML2 is required for MAC, the integrity of the label must be provided with SML2, and the function (possibly a cryptographic algorithm) binding the label to the data must also be SML2. Other implementation concerns include making the labeling non-bypassable and fail-safe.
- *Discretionary Access Control (DAC)* is different from MAC in that the owner of the data to be accessed (versus the machine) can choose who can and cannot be authorized access to the data. For SML1, this is comparable to setting UNIX permission bits (owner/group/world) to grant access. For SML2 and 3, using ACLs further refines the mechanism. ACLs can be more specific to allow certain identities access to information, (e.g. specific users within a group can be granted access). Again, DAC mechanisms should be non-bypassable (only “changeable” by the owner of the data), fail-safe, and possess the same SML level of integrity associated with the level of DAC required.
- *Certificates* are described under I&A (Table 4-12).
- *Personnel security* is described under I&A (Table 4-12).

**Table 4-13 Access Control Mechanisms**

	Anti-Tamper	Mandatory Access Control	Discretionary Access Control	Certificates	Personnel Security
<b>SML1</b>	level 1 or 2 [4]	Not applicable	Comparable to Unix permission bits	bind w/SML1 cryptographic algorithm	Commercial hiring practices
<b>SML2</b>	level 3 or 4 [4]	Labels bound to data having integrity and binding function both at the SML2 level	ACLs	bind w/SML2 cryptographic algorithm	Equivalent of Secret clearance
<b>SML3</b>	level 4 or better [4]	Labels bound to data having integrity and binding function both at the SML3 level	ACLs	bind w/SML3 cryptographic algorithm	Equivalent of Top Secret clearance

### 4.5.3.7 Mechanisms Supporting Accountability

Accountability can be considered a special case of non-repudiation. The accountability security service is basically holding any network entity responsible for its actions on that network. Mechanisms, that can be used to provide the security service of accountability are shown in Table 4-14, and discussed below.

- When implementing the *audit* mechanism, the following components should be considered.
  - What is being audited and relevant events that are detected.
  - How the audit (detected) data is protected, analyzed, and reported.



- What the reaction strategy is to the audit data analysis and reporting.

These components should be considered for each SML level, and in SML2 and 3, should be detailed in a plan. As with all mechanisms, consideration should be given to noncircumvention or “non-bypassability” and the effects of failure.

- *Intrusion detection* is still in relative infancy. Intrusion detection is that mechanism that monitors a network and detects either (1) known attacks being mounted against the system or (2) differences in a profiled use of the system. Several aspects associated with an intrusion detection mechanism for examples, whether it is static [SML1] set up to filter only on known attacks and profiles); dynamic [SML2] set up to filter on known attacks and profiles but updateable perhaps through software downloads), or dynamically adaptable [SML3] this adds the aspect of “artificial intelligence” in which the system learns new profiles based on usage). Depending on the SML level, a reaction mechanism to a detected intrusion must be either informally (SML1) or formally (SML2 and 3) detailed and implemented.
- *I&A* is described under I&A (Table 4-12).

**Table 4-14 Accountability Mechanisms**

	<b>Audit</b>	<b>Intrusion Detection</b>	<b>I&amp;A</b>
<b>SML1</b>	informal reaction mechanism	static system with informal reaction mechanism	see I&A table for SML1
<b>SML2</b>	formal reaction plan and strategy	dynamic system with formal reaction mechanism	see I&A table for SML2
<b>SML3</b>	formal reaction plan and strategy	dynamic, adaptive system with formal reaction mechanism	see I&A table for SML3

### 4.5.3.8 Mechanisms Supporting Non-Repudiation

The security service of non-repudiation provides a method by which the sender of data is provided with proof of delivery and the recipient is assured of the sender’s identity, so that neither can later deny processing the data. It is quantified in Table 4-15 and can be obtained by a combination of these mechanisms as appropriate for the applications:

- *Signature* is used to digitally sign data in such a way that only the sender and receiver could have respectively sent and received the message. The sender signs the original data to prove that it was sent. The receiver signs a receipt as proof of receipt of the original data. Validation of these signatures is always required.
- *Trusted Third Party* is used to prearrange a method by which a third party may receive the information from the sender and transmit/send it to the receiver in a way that ensures that the sender and receiver are confident that they are communicating with the correct party.
- *Accountability* is described under Accountability (Table 4-14)
- *I&A* is described under the I&A (Table 4-12).
- *Archive* is the ability to store data so that it can be recovered if necessary.

**Table 4-15 Non-Repudiation Mechanisms**

	<b>Signature</b>	<b>Trusted Third Party</b>	<b>Accountability</b>	<b>I&amp;A</b>	<b>Archive</b>
<b>SML1</b>	sign with SML1 cryptographic algorithm	see I&A Table for SML1 Personnel Security	see Accountability table for SML1	see I&A table for SML1	informal archival plan, user backs up own key or data
<b>SML2</b>	sign with SML2 cryptographic algorithm	see I&A Table for SML2 Personnel Security	see Accountability table for SML2	see I&A table for SML2	formal archival plan, central back- ups
<b>SML3</b>	sign with SML3 cryptographic algorithm	see I&A Table for SML3 Personnel Security	see Accountability table for SML3	see I&A table for SML3	formal archival plan, central, offsite back-ups

## 4.5.4 Level of Assurance

The discussion addressing the need for an overall system security solution view of strength of mechanism is also relevant for the level of assurance. Again, while an underlying methodology is offered, a real solution can only be deemed effective after a detailed analysis activity that considers the specific operational and threat situations and the system context for the solution.

Assurance is the measure of confidence in claims made and that the security features and architecture of an automated information system appropriately mediate access and enforce the security policy. The assurance measures listed here are from the Common Criteria [6].

The Common Criteria provide assurance through active investigation. Active investigation is an evaluation of the actual product or system to determine its actual security properties. The Common Criteria philosophy assumes that greater assurance results come from greater evaluation effort in scope, depth, and rigor. This leads to the seven EALs (EAL1 through EAL 7) described below:

- *EAL 1, Functionally Tested*, is applicable where some confidence in correct operation is required, but the threats to security are not viewed as serious. It is of value where independent assurance is required to support the contention that due care has been exercised with respect to the protection. An example is the protection of personal information.
- *EAL 2, Structurally Tested*, requires the cooperation of the developer in terms of the delivery of design information and test results, but should not demand more effort (or substantially increased cost or time) than is consistent with good commercial practice. It is applicable where a low to moderate level of independently assured security is required in the absence of an available development record. An example is securing legacy systems, or where access to the developer is limited.
- *EAL 3, Methodically Tested and Checked*, permits a conscientious developer to gain maximum assurance from positive security engineering at the design stage without substantial alteration of existing sound development practices. It is applicable where a moderate level of independently assured security is required.

- *EAL 4, Methodically Designed, Tested and Reviewed*, permits a developer to gain maximum assurance from positive security engineering based on good commercial development practices which, though rigorous, do not require substantial specialist knowledge, skills, and other resources. This is the highest level at which it is likely to be economically feasible to retrofit to an existing product line. It is applicable in those circumstances where a moderate to high level of independently assured security in conventional products is required, and where developers or users are prepared to incur additional security-specific engineering costs.
- *EAL 5, Semi-Formally Designed and Tested*, permits a developer to gain maximum assurance from security engineering based on rigorous commercial development practices supported by moderate application of specialized security engineering techniques. It is applicable where a high level of independently assured security in a planned development is required along with a rigorous development approach.
- *EAL 6, Semi-Formally Verified Design and Tested*, permits developers to gain high assurance from application of security engineering techniques to a rigorous development environment to protect high value assets against significant risks. It is applicable to the development of security products to be used in high-risk situations.
- *EAL 7, Formally Verified Design and Tested*, is applicable to the development of products to be used in extremely high risk situations and/or where the high value of the assets justifies the higher costs. Realistically, it is limited to products with tightly focused functionality that is amenable to extensive formal analysis.

The assurance levels are a composition of the following assurance classes: Configuration Management, Delivery and Operation, Development, Guidance Documents, Life-Cycle Support, Tests, and Vulnerability Assessments. These classes incorporate the concepts of correct implementation, “non-bypassable” mechanisms, failure to a secure state, secure startup, and others.

In addition to those addressed in the Common Criteria, there are other assurance tasks that the Common Criteria do not discuss, including Failure Analysis and Test, TEMPEST Analysis and Test, and Tamper Analysis and Test. If these apply to a particular product or system, then they should be added to the requirements of the appropriate EALs.

## 4.5.5 A Worked Example

### Assumptions

- Security evaluation is a necessary part of solution development.
- A trained ISSE (or equivalent) is the strategy consumer.

The methodology for correct employment of the robustness strategy is as follows.

- The responsible customer party knows, and has appropriately documented, mission objectives, concept of operation, value of information to be protected, threat/environment context, and security policy.

- A solution is then engineered (per IATF Chapters 5 through 9 [Defend the Network and Infrastructure, Defend the Enclave Boundary/External Connections, Defend the Computing Environment, Supporting Infrastructures, and Information Assurance for the Tactical Environment], providing guidance on the security mechanisms required.
- Risk factors (e.g., degree of damage if security policy is violated, threat environment) will be used to help determine the appropriate strength and associated level of assurance for each mechanism from the set of security service tables. The risk addressed is the residual risk, not the overall (or initial) risk, what remains after other countermeasures have been applied, and what would be the target of doctrine if additional security measures were not taken. For example, a system-high workstation in a “secure” office setting has a different residual risk from that same workstation operating in a public environment.
- Working with an ISSE, the customer will then select COTS/GOTS product(s) providing the necessary strength and assurance.
- Evaluate the system and highlight the residual risk.

### 4.5.5.1 Example One

The following illustration uses an abbreviated example of the media protection portion of the IATF Remote Access (Section 6.2), Secret Dial-in Case, to demonstrate how the robustness strategy would typically be used in conjunction with other guidance sections of the IATF. No attempt was made to consider an actual customer’s needs or an actual recommended solution.

In this example, the customer will be processing Secret data at a Continental U.S. (CONUS) site (perhaps in a work-at-home or temporary duty (TDY) situation) on a remote access dial-in system. The customer is required to protect this data and feels the threat to the data is primarily from *adversaries*:

- *Who have minimal resources at their disposal* (i.e., they have enough money or contacts so that they can get someone to steal the laptop out of a house/hotel room)
- *Who are willing to take significant risk* (i.e., if the person gets caught stealing, the adversaries are willing to be prosecuted; or know that if the thief gets caught the theft will not be traced back to them).

For this example, a media encryptor is recommended to provide confidentiality of the customer’s Secret data on the hard drive of the remote computer. Because the data is Secret, according to the current classification manual, compromise of that data would *cause serious damage to the security of the United States*. Based on the situation described here, the customer, in conjunction with the ISSE, determines that the value of his information is at the *V4* level (violation of the information protection policy would cause serious damage to the security, safety, financial posture, and/or infrastructure of the organization), and the perceived threat is at the *T3* level (adversary with minimal resources who is willing to take significant risk). Using the Degree of Robustness Table presented in Table 4-7 (as depicted in Table 4-16), the minimum SML and EAL recommended is *SML2* and *EAL3* based on the threat and information levels.

**Table 4-16 Example Depicting Use of Degree of Robustness Table**

Information Value	Threat Levels						
	T1	T2	T3	T4	T5	T6	T7
<b>V1</b>	SML1 EAL1	SML1 EAL1	SML1 EAL1	SML1 EAL2	SML1 EAL2	SML1 EAL2	SML1 EAL2
<b>V2</b>	SML1 EAL1	SML1 EAL1	SML1 EAL1	SML2 EAL2	SML2 EAL2	SML2 EAL3	SML2 EAL3
<b>V3</b>	SML1 EAL1	SML1 EAL2	SML1 EAL2	SML2 EAL3	SML2 EAL3	SML2 EAL4	SML2 EAL4
<b>V4</b>	SML2 EAL1	SML2 EAL2	<b>SML2</b> <b>EAL3</b>	SML3 EAL4	SML3 EAL5	SML3 EAL6	SML3 EAL6
<b>V5</b>	SML2 EAL1	SML2 EAL3	SML3 EAL4	SML3 EAL5	SML3 EAL6	SML3 EAL6	SML3 EAL7

For our example, and as a result of applying the IATF guidance, the ISSE together with the customer determined that confidentiality and security management services are recommended. The user of the remote access dial-in system will want to keep the Secret data on the laptop inaccessible while in storage. Not only does the data need to be encrypted on the media, but also the system needs to be operated in a secure manner; furthermore, the issue of how to recover the data if it is compromised needs to be addressed. The ISSE and customer together decide that media encryption will be one mechanism used. From the discussions above, a media encryptor of strength SML2 should be considered.

Once the security service has been selected, confidentiality in this case, then the mechanism should be chosen from the columns of that table. In this case, the mechanism chosen is cryptographic algorithm. This mechanism has been chosen because it was the cheapest, simplest, most practical to implement. Physical security was not chosen because it was impossible to uniformly apply, in a timely manner, at different remote sites without knowing all the sites in advance. Technical security was not chosen because of the wide variety of COTS laptops, which are currently not built with technical security countermeasures. Using the Confidentiality Mechanisms Table, as depicted in Table 4-17, the implementation should look for a cryptographic algorithm capability with an effective key length of 80+ bits, supported by a KMI/PKI providing the strength under category “Y” as further described in Chapter 8-1 (KMI/PKI).

**Table 4-17 Example Depicting Use of Confidentiality Mechanisms Table**

	Cryptographic Algorithm		Physical Security	Technical Security		Anonymity	
	Effective Key Length	Key Management		Anti-Tamper	TEMPEST	TRANSEC	Cover
<b>SML1</b>	40+ bits symmetric key length, 80+ exponent 512+modulus public key length	SMI Cat X, 80+exponent 512+ modulus public key length, 80+ hash key length	Comparable to [7]	level 1 or 2 [4]	comply with applicable EMI/EMC FCC standards or portions of [8]	low power unit	TBD
<b>SML2</b>	80+ bits symmetric key length, 160+ exponent 1024+ modulus public key length	SMI Cat, 160+ exponent 1024+ modulus public key length, 160+ hash key length	Comparable to [7]	level 3 or 4 [4]	[8]	commercial spread spectrum signal techniques	TBD
<b>SML3</b>	Because of to the complicated nature of this level, please consult with a qualified ISSE. <sup>6</sup>	SMI Cat Z, also consult with a qualified NSA ISSE. <sup>6</sup>	Comparable to [7]	level 4 or better [4]	[8]	cryptographic spread spectrum signal techniques	TBD

Because the remote access dial-in users will not have direct access to their system administrator or support services, the customer and ISSE found that the need for the security management mechanisms of training and secure operations was of paramount importance and should be supplied at the SML3 level. Similarly, because of the “remote” use of the system, they thought that compromise might be more likely; and, therefore, the compromise recovery mechanism was also of paramount importance and should be addressed at the SML3 level. Further, because of the value of the information and the threat to the information, the components should be characterized as methodically tested and checked, consistent with the Common Criteria EAL3. (Note that this depicts a situation where the initial SML and EAL recommendations from the strategy were considered inadequate, and were thus increased, presumably based on a detailed analysis of the situation.) Table 4-18 depicts how the Security Management Mechanisms table would typically be used.

Note that when using the tables in this section, not all columns must be used, and various SML levels may be employed as needed for the specific mechanism under question. In the media encryption example, it may be determined that security management mechanisms are of paramount importance; therefore, *SML3* will be chosen while confidentiality mechanisms to protect data may be adequate with a *SML2* cryptographic algorithm.

<sup>6</sup> DoD users should consult with a National Security Agency ISSE. Other government users are directed to contact an ISSE at the National Institute of Standards and Technology for guidance in this area. Non-government users should consult with a qualified ISSE, or equivalent representative within their organization.

**Table 4-18 Example Depicting Use of Security Management Mechanisms Table**

	<b>Compro- mise Recovery</b>	<b>System Admin- istration</b>	<b>Training</b>	<b>OPSEC</b>	<b>Trusted Distribution</b>	<b>Secure Operations</b>	<b>Mechanism Manage-ment</b>
<b>SML1</b>	informal plan	See Ch. 4 countermeasures	training available at user discretion	Implement OPSEC at user's discretion	direct vendor purchase	informal plan of operation	procedural, user's discretion
<b>SML2</b>	detailed plan that is reviewed and approved	See Ch. 4 countermeasures	formal training plan	OPSEC training required, implement at user's discretion	certificate of authenticity, virus scan, validation	formal plan of operation	procedural, reminders, user's discretion
<b>SML3</b>	detailed plan that is reviewed and approved	See Ch. 4 countermeasures	Knowledge/skill certification required	OPSEC training required, implement OPSEC required	protective packaging, checksums, validation suite	detailed, formal plan of operation	automated support

### 4.5.5.2 Example Two

A second example of the use of the strategy is where a Sensitive Compartmented Information Facility (SCIF) is employed as physical protection. Very different security mechanisms would likely be chosen to protect the information. If a DoD system is processing Top Secret data (V5), and the threat is very high (T6), one would normally apply rigorous SML and EAL levels for this solution. However, because the SCIF is used (and there is no connectivity outside the SCIF), the confidentiality requirement is mostly satisfied by physical security at SML3 level. The access control requirement may also be satisfied by personnel security at SML3 level. The residual risk in the areas of confidentiality and access control may be mitigated by additional mechanisms at the SML1 level. This example shows the importance of layering security mechanisms to reduce risk.

### 4.5.5.3 Example Three

A third example involves a corporation with a large Intranet that processes only unclassified data. The corporation has stringent legal requirements to protect its data from unauthorized access or modification. The corporation maintains a large heterogeneous network with Internet access protected by firewalls. All data requiring legal protection is maintained in isolated subnets and is not available to authorized users via the network. Offline stand-alone access is required to view the protected data. The security objective is to upgrade the network to allow the protected data to be securely accessible by all authorized users. Although the data being processed is unclassified but it must be protected from unauthorized access. Using the applicable capability maturity model, a Capability Level 2 or equivalent is recommended. Taking all this into consideration, the customer along with the ISSE determined that the information is at the V3 level (violation of the information protection policy would cause some damage to the security safety, financial posture, and/or infrastructure of the organization), and the perceived

threat is at the *T4* level (sophisticated hackers, international corporations). Using the Degree of Robustness Table presented in Table 4-7 (as depicted in Table 4-19), the minimum SML and EAL recommended is SML2 and EAL3 based on the threat and information levels.

**Table 4-19 Example Depicting Use of Degree of Robustness Table**

Information Value	Threat Levels						
	T1	T2	T3	T4	T5	T6	T7
<b>V1</b>	SML1 EAL1	SML1 EAL1	SML1 EAL1	SML1 EAL2	SML1 EAL2	SML1 EAL2	SML1 EAL2
<b>V2</b>	SML1 EAL1	SML1 EAL1	SML1 EAL1	SML2 EAL2	SML2 EAL2	SML2 EAL3	SML2 EAL3
<b>V3</b>	SML1 EAL1	SML1 EAL2	SML1 EAL2	SML2 EAL3	SML2 EAL3	SML2 EAL4	SML2 EAL4
<b>V4</b>	SML2 EAL1	SML2 EAL2	SML2 EAL3	SML3 EAL4	SML3 EAL5	SML3 EAL5	SML3 EAL6
<b>V5</b>	SML2 EAL2	SML2 EAL3	SML3 EAL4	SML3 EAL5	SML3 EAL6	SML3 EAL6	SML3 EAL7

In examining at the corporation's security objectives, the customer and ISSE determined that *access control* to the sensitive data and *confidentiality* of the data as it transits over the intranet are the security services required. The mechanism(s) for implementation need to operate on both Windows NT and HP UNIX platforms.

The confidentiality mechanisms for the SML2 category recommend a minimum 80+ bit symmetric key length, 160+ exponent 1024+ modulus public key length. The firewall key scheme includes ISAKMP/OAKLEY with Data Encryption Standard (DES) or 3DES capability. 3DES is the scheme being evoked. The I&A mechanisms for the SML2 category recommend a system ID and a password with minimum character lengths. The corporation implements user IDs that are a minimum of six characters and eight characters for passwords with an alphanumeric mix. However, this is an internal Intranet, no security services for integrity, availability, and non-repudiation are considered necessary.

Each server requiring protection will have an individual firewall installed with the rules base requiring positive user identification and authentication before access is allowed. Initially, this process will be accomplished using use of user IDs and passwords; however, it will migrate to a PKI certificate based capability. Confidentiality will be provided by the Virtual Private Network (VPN) capability resident to the firewall product. Client VPN software will be installed on each client machine enforcing the connection and VPN rules to the protected servers (if the client VPN is disabled, no connection is allowed to a protected server).

The following security mechanisms are employed.

- Fronting each server that contains protected data with a firewall.
- Invoking VPNs between client machines and the server and printers (using 3DES algorithm).



- Implementing user identification and authentication using the VPN user ID and password.
- Implementing the firewall rule base to allow access by authorized only users from authorized workstations.
- Consideration is being given to replacing the VPN only client with a client that provides the VPN capability and extends the firewall policies to the user's desktop.

## 4.5.6 Robustness Strategy Evolution

Although “robustness” is now an inherent part of the IATF, it is a relatively new term in the IA lexicon and is not clearly seen as a unifying successor to a variety of similar existing concepts such as completeness, assurance, and accreditation.

The security mechanism tables provide guidance at three strength levels to support a variety of security services. At another level of table refinement, security functions would appear, each of which would implement a particular mechanism. For example, each specific cryptographic algorithm would be a security function to implement a cryptographic algorithm mechanism in support of, say, a confidentiality security service. Many security functions that implement each mechanism.

To compare and contrast these functions, there needs to be a way to cost the relative strengths. This effort would require developing cost metrics for each security service. Although functional specifications might be a relatively modest enhancement, the development of multiple costing schemes is likely to be a monumental effort. This level of refinement, which would enable uniform comparison of the protection provided by security mechanisms, is the goal of the strategy.

The IATF layered approach to security means that a variety of services and mechanisms may be necessary to achieve the necessary protection. A broader view needs to be developed, looking across all needed services and the mechanisms proposed to provide those services. The residual risk to a system product needs to be addressed based on the environment in which it is implemented.

In addition to the above concerns, and because threat environments and security technologies are changing continually, the guidance provided is subject to frequent revision. To the extent possible, all mechanism recommendations should be by indirect references to formally endorsed documents. In cases where this is not possible, periodic revision and trained ISSE application is the best way to ensure that guidance is current.

## 4.5.7 Real World Applications

In the real world, it quickly becomes too complicated and impractical to determine layered solution approaches and describe, offer, support, and implement them for more than a small number of different robustness levels. The threat levels and information value levels described earlier simply yield too many combinations of SML and EAL levels, as shown in Table 4-7. The Office of Secretary of Defense (OSD) Information Assurance guidance and policy for the DoD's Global Information Grid (GIG) divides robustness into three levels, a more practical approach.

The OSD GIG policy uses an implementation approach for robustness that makes conclusions based on real-world conditions (see Appendix E, OSD IA Policy Robustness Levels).

### 4.5.7.1 Future Work

The following areas need further attention.

- The Network Rating Model/Methodology also addresses “goodness.” How can that effort be incorporated into the strategy?
- The issue of composition of metrics needs to be addressed in the framework of layered security.
- There is a need to ensure that terminology used in the strategy is definitive and consistent with the remainder of the IATF.
- The current approach to security is considered non-scalable which means the process used for small systems may not be appropriate for large systems. This is also known as the composibility problem and the layering problem. How can the robustness strategy help address this issue?
- The mechanism tables need to be reviewed for non-uniform level of detail and non-quantifiable entries.
- Update the strategy to incorporate Common Criteria “language” throughout, rather than just describing the EALs.
- Consider the effect of recommended robustness on return on investment to the customer.

## 4.6 Interoperability Framework

Users continue to become more dependent on information systems, fostering a driving need for connectivity and interoperability at the application level. As information and telecommunication systems are being introduced and updated, interoperability of these systems is a major concern of the organizations that use them. When these systems have to be secure, efficient interoperability becomes more difficult to achieve and manage. This section of the Framework provides a high level strategy for dealing with interoperability at the architecture and technology levels. Later releases of this Framework will address the issue of interoperability comprehensively, making users aware of options and tradeoffs, and providing guidance addressing with this important challenge.

### 4.6.1 Major Elements of Interoperability

This section identifies numerous elements that must be addressed in order to achieve interoperability. To achieve interoperability, it is typically necessary to deal with all of them. These elements and the issues associated with them are discussed below.

- *Architecture.* A first step in achieving interoperability is an agreement on the nature of the security services, the type of security mechanism(s) to be used, and their allocation

to functional components (e.g., enclave boundary interfaces, end user terminals of the architecture, and the layers at which security mechanisms are applied).

- *Security Protocols.* It is fundamental that systems use compatible communications protocols to achieve user-to-user connectivity. When this connectivity must be secure, several security elements associated with security protocols also must be considered. These elements include security services, cryptographic algorithms (with modes and bit lengths), synchronization techniques, and key exchange techniques. If options are permitted, common provisions are also needed for algorithm selection and broader security option negotiation. Typically, security protocol designers deal with these elements.
- *Product Compliance to Standards.* Another element needed for interoperability stems from the assurance that products used to implement a network security solution actually comply with the standards they claim to support. There are a number of initiatives with the commercial sector and in government that will verify compliance, as discussed below.
- *Interoperable KMI/PKI Support.* The services and techniques used to provide KMI/PKI are another element needed to achieve interoperability. This includes key and certificate formats, token mechanisms, cross certification (to facilitate communication across KMI/PKI security domains), directory systems, and compromise recovery capabilities. These considerations are discussed further in Section 4.7 (Key Management Infrastructure/Public Key Infrastructure Considerations).
- *Security Policy Agreement.* Beyond all of the technical issues needing to be addressed to allow interoperability is a fundamental issue of organizational security policies that establish the ground rules for permitting interoperability. It is necessary that the network or system “owners” determine what minimum protection mechanisms and assurances (perhaps for particular types of data or destinations) are needed before they would be willing to allow users from other networks or systems to communicate or interact with users of their resources and information. Because this important topic is beyond the scope of this document, it is assumed in this Framework that organizations wishing to interoperate have resolved any incompatibilities in organizational security policy and that the only barriers are technical or economic.

## 4.6.2 Challenges for Interoperability

When formulating an IA solution, the following potential impediments tend to act as obstacles to achieving interoperability:

- Backward compatibility with legacy systems that do not use accepted standards, and lack negotiation mechanisms needed to interoperate with newer standards-based implementations (even if backward-compatible protocols and modes are available).
- Security solutions lagging the rapid pace of evolving information technologies, often making security an adjunct capability.
- Evolution of standards or lack of standards accepted by either the user community or commercial product marketplace.
- Defacto proprietary standards or closed systems.

- Lack of an accepted source of testing to verify that products implementing standards do so correctly and that sufficient options of the standards are implemented, so users can be assured that the resultant products are in actuality, interoperable.

The challenge is to recognize and accommodate these obstacles, yet still find a way to achieve the interoperability needed by our customers.

### 4.6.3 Interoperability Strategy

At this point in the IATF, it is appropriate to establish a basic, high level strategy for dealing with interoperability. This strategy can be characterized by specific efforts focused on the following.

- Fostering standards for secure applications and communications protection that are based on open architectures.
- Supporting security negotiation protocol standards that allow users to have varying policies and provide a vehicle to negotiate elements of interoperability.
- Developing a migration strategy from the interim solutions to open standards in environments where emerging technology dominates and users accept interim solutions that are not standards based.
- Defining initial interoperability standards, and influencing and migrating to a standards-based approach where gaps exist.

A major issue still remains. It is imperative to ensure that products and system components correctly implement these standards and options so interoperability is actually realized. A number of initiatives within the government and private sectors exist to address this issue. These include the following.

- *Automotive Network eXchange® (ANX)*—The automotive industry has recognized the importance of interoperability for the transport of trading partner electronic information. The ANX network service is positioning to provide automotive trading partners with a single, secure network for electronic commerce and data transfer—replacing the complex, redundant, and costly multiple connections that exist throughout the automotive supply chain.
- *International Computer Security Association (ICSA)*—The ICSA promotes the open exchange of information between security product developers and security service providers. ICSA acts as an independent, third party that offers a number of initiatives, including a product certification program. The ICSA certification develops criteria by which industry wide categories of products are tested. The ICSA certifies products on an annual basis and spot-checks for compliance throughout the year against the latest version of each product. Using this process, buyers of ICSA-certified products can be assured of getting the most secure products available at the time.
- *National Information Assurance Partnership (NIAP)*—The NIAP is a joint industry/government initiative, lead by the National Institute of Standards (NIST) and NSA to establish commercial testing laboratories where industry product providers can have security products tested to verify their performance against vendor claims. As with the

ICSA initiatives, a natural fallout of this testing will be assurance that users can be assured that those products advertising compliance with standards will indeed be interoperable.

These activities, and a number of others similar to them, will help product and system providers to deliver solutions that will support the interoperability needs of their broad customer base.

The interoperability strategy presented in this section is embodied throughout this IATF. In a later release of the IATF document, a more detailed treatment of specific issues impacting interoperability will be included in subsequent sections. Specifically, Chapters 5 through 9 (Defend the Network and Infrastructure, Defend the Enclave Boundary/External Connections, Defend the Computing Environment, Supporting Infrastructures, and Information Assurance for the Tactical Environment) of the IATF will include discussions of interoperability issues specific to each of the user requirement categories. These will include interoperability concerns or needs reflected in the captured requirements, the technology assessments (to identify the degree to which the available solutions deal with interoperability issues), and the recommendations (that deal with selection of architectures and protocols that achieve the needed interoperability). Chapter 8 (Supporting Infrastructures) will deal specifically with interoperability issues associated with KMI/PKI.

## 4.7 Key Management Infrastructure/ Public Key Infrastructure Considerations

A KMI/PKI capability is needed to support most technical security countermeasures. This section provides a high level discussion of the role of, and features associated with, a KMI/PKI. Detailed guidance for the architecture of KMI/PKI can be found in Chapter 8 (Supporting Infrastructures) of the Framework.

### 4.7.1 KMI/PKI Overview

The KMI/PKI process generates, distributes, and manages security credentials. It can be considered as a set of interrelated activities providing security services that are needed to enable the Framework's security solutions presented in Chapters 5, 6, 7, and 9 (Defend the Network and Infrastructure, Defend the Enclave Boundary/External Connections, Defend the Computing Environment, Supporting Infrastructures, and Information Assurance for the Tactical Environment). KMI/PKI is a unique user requirement category in the Framework because it does not directly satisfy a user's security requirements; rather, it facilitates the use of security building blocks that are needed by other security mechanisms.

Current KMI/PKI implementations consist of numerous stovepipe infrastructures that support different user solutions. These are run by various organizations, even though the end user may need support from several stovepipes for a single application. A critical aspect of a complete system approach to any network security solution needs to include a KMI/PKI architecture that provides effective and efficient operations while maintaining the requisite security features and assurances.

A KMI/PKI architecture is heavily dependent on the specific applications it supports. For example, a VPN provides an encrypted pipe between two enclaves. The KMI/PKI provides keys and certificates to the cryptographic devices that provide authentication and encryption used to establish and maintain the pipe. KMI/PKI could also provide additional services, including data recovery and a directory to provide access to users' public certificates.

A second area where KMI/PKI differs from other solutions in the Framework is that its security is distributed throughout a number of separate elements. These elements require extensive security (e.g., encryption, certificate management, compromise recovery), among themselves to protect the user's key or certificate. Because of the repercussions of a successful attack against the KMI/PKI, internal infrastructure security requirements are often more stringent than user services security. There are also unique requirements also exist on the infrastructure (e.g., policy management), and the level of assurance for the KMI/PKI services is often higher.

## 4.7.2 KMI/PKI Operational Services

Section 8.1 (KMI/PKI) addresses four operational services supported by the KMI/PKI. These services support different user applications and consequently employ different (but related) mechanisms and have unique security requirements. The first user service is symmetric key generation and distribution. This is still the primary key management mechanism within the classified community.

The second service, PKI, addresses both digital signature (for authentication and integrity) and key agreement with its associated certificate management. This is the primary key management mechanism within the commercial community.

The third service, directory service, is used to provide access to the public information required with PKI such as the public certificate, the related infrastructure certificates, and the compromised key information. Directory services can be provided either by a global set of distributed directories (e.g., X.509 Defense Message System [DMS] directories), or by an on-line repository at a single site. Although directories can be used for other things, they are normally very closely coupled with PKI.

The final service is managing the infrastructure itself. The distributed nature of the infrastructure places additional functional and procedural requirements on the KMI/PKI and the sensitivity of the application places additional security requirements on the KMI/PKI. The internal structure of the infrastructure varies with the application it supports.

## 4.7.3 KMI/PKI Processes

The KMI/PKI consists of a numerous processes that all have to work together correctly for a user security service to be truly secure. Each of these processes is necessary at some level in all KMI/PKI architectures. The processes include the following.

- *Registration*—Authorizing people-making decisions about the validity of user actions.
- *Ordering*—Requesting the KMI/PKI to provide a user either a key or a certificate.

- *Key Generation*—Generating of the symmetric or asymmetric key by an infrastructure element.
- *Certificate Generation*—Binding the user information and the asymmetric key into a certificate.
- *Distribution*—Providing the keys and certificates to the user in a secure, authenticated manner.
- *Accounting*—Tracking the location and status of keys and certificates.
- *Compromise Recovery*—Removing invalid keys and certificates from the system in an authenticated manner.
- *Re-key*—Periodically replacing keys and certificates in a secure, authenticated manner.
- *Destruction*—Destructing the Secret key when it is no longer valid.
- *Data Recovery*—Being able to recover encrypted information without direct access to the original key.
- *Administration*—Running the infrastructure.
- *Value-added PKI Processes*—Supporting optional value-added processes including archive, time-stamp, and notary services, (PKIs only).

The complete set of KMI/PKI processes are usually distributed to several elements performing independent tasks, requiring extensive coordination and security processing between elements. For most processes, numerous ways exist to implement the services based on the application supported, the security required, and the cost (e.g., money, people, and performance) the user is willing to pay. Each process contributes to the overall security of the KMI/PKI and has various forms of attacks and countermeasures.

## References

1. Jeff Humphrey and Bruce Gabrielson, Ph.D. “Phreakers, Trashers, and Hackers,” presented at AFSEA INFOSEC Engineering Course, 1995, Burke, VA  
<http://blackmagic.com/ses/bruceg/hackers.html>
2. Root Shell at <http://www.rootshell.com>
3. Coast Security Pages at <http://www.cs.purdue.edu/coast/intrusion-detection/>
4. FIPS PUB 140-1, National Institute of Standards and Technology, Security Requirements for Cryptographic Modules, <http://www.itl.nist.gov/div897/pubs/fip140-1.htm>
5. NSTISSI No. 4009, National INFOSEC Glossary.
6. Common Criteria for Information Technology Security Evaluation, CCIB-98 (ISO/IEC 15408), Version 2.0, 1998, <http://csrc.nist.gov/cc/>.
7. DoD Reg. 5200.1-R, Information Security Program, 1997.
8. NSTISSAM TEMPEST/1-92, Compromising Emanations Laboratory Test Requirements, Electromagnetics, 1992.
9. Laing, Alan, “DoD PKI Level of Protection and The Appropriateness of Proposed Solutions for Various Applications”, Draft ,1998.
10. National Security Agency Specification for General Functional Security Requirements for a Telecommunications System (FSRS), 1989.
11. Information Systems Security Engineering Handbook, Release 1.0, 28 February 1994.
12. Security Management Infrastructure (SMI) Task 1 Team, Threat and Vulnerability Model for Information Security, 1997.

## Additional References

- a. NSA/CSS Dir. No. 120-1, NSA/CSS Operations Security Program, 1990.
- b. National Security Agency Specification for Unified INFOSEC Criteria, 1991.
- c. 16 Warwick Ford, *Computer Communications Security*, Prentice Hall PTR, Englewood Cliffs, NJ, 07632, 1994.



## Chapter 5

# Defend the Network and Infrastructure

Networks provide a transport mechanism for user traffic and for the availability of user information. Networks and their supporting infrastructures must protect against denial of service attacks that could prevent user information from being transmitted. The supporting infrastructure is the management systems and any other systems that support network operation.

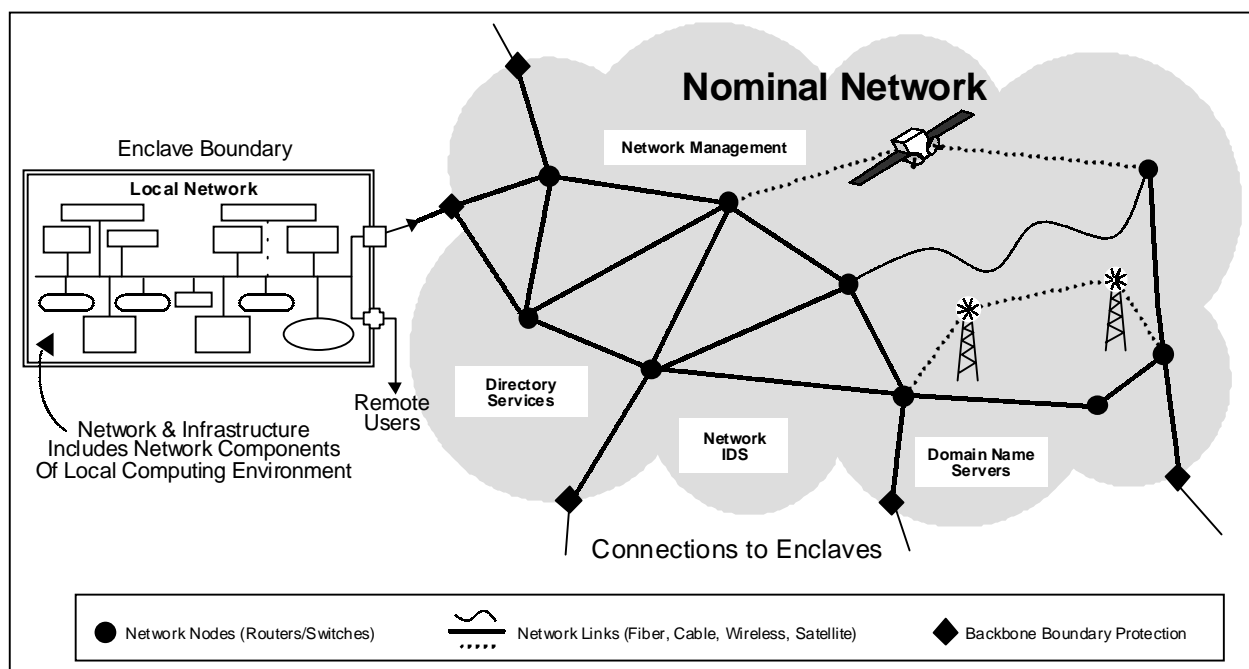
The network supports three distinct types of traffic: user, control, and management. User traffic is simply the information that users' are transmitting over the network. Networks have the responsibility to provide separation of user traffic. Isolation of individual user connections must be maintained to ensure reliable delivery of information. Additionally, confidentiality services *may* be provided by the network, either by government encryptors for classified traffic or through commercial encryption embedded in network components for unclassified traffic.

Control traffic is any information transferred between network components that is necessary for establishing user connections. Control traffic provided by a signaling protocol, such as Signaling System 7 (SS7), includes addressing, routing information and signaling. Proper addressing by the network infrastructure is essential for user traffic to be directed to the intended destination. Routing information must be protected to ensure that the user information will be properly transferred and that the path that user information takes is not manipulated. Similarly, signaling must be protected to ensure user connections are established properly.

The third type of network traffic, management traffic, is any information that configures network components or information initiated from a network component that informs the network infrastructure on status of the network component. Management protocols include Simple Network Management Protocol (SNMP), Common Management Information Protocol (CMIP), Hypertext Transfer Protocol (HTTP), rlogin and telnet command line interfaces, or other proprietary management protocols. Network management traffic protection is essential to ensure that network components are not modified by unauthorized users. If management of a network component is compromised, that component can be configured to perform any function the attacker wishes. Simply being able to view configuration information on a network component may give an attacker knowledge of network connections, addressing schemes, or other potentially sensitive information. Figure 5-1 illustrates the network and infrastructure in the high level Defense Information Infrastructure (DII) context. Some of the networks illustrated are controlled by government organizations, while others are controlled by commercial entities such as the Public Switched Telephone Network (PSTN) and the Internet.

Today, commercial carriers provide over 95% of all the transmission service for all communications, the Federal Government and industry. In addition, most of the large civil Government networks provided by General Services Administration (GSA), Federal Aviation Administration (FAA), Department of Transportation, etc., outsource the management of their networks. In light of this reliance on commercial control networks, all organizations should adopt a two-pronged approach—starting at the highest level—to defend their networks. First, organizations

should ensure they have established clear Service Level Agreements (SLA) with their commercial carrier that specify metrics for reliability, priority, and access control. Commercial carriers view network security as a business issue. Therefore, they will not simply add security features. For them a business case must be made; the customer must ask for these services. Secondly, organizations should also recognize that during transmission, their data may be essentially unprotected. It is incumbent upon the *owner* of the information to implement security services, such as encryption for confidentiality, at the user level. Historically few organizations outside the Department of Defense (DoD) and the Intelligence Community (IC)—have developed strategies and encrypted data sent over commercial lines. In the past few years, however, services such as Pretty Good Privacy (PGP) have grown in use by government and industry organizations.



**Figure 5-1 Defend the Network & Infrastructure**

The general IA strategy for defending the network and infrastructure is to use approved Wide Area Networks (WANs) to transport classified data among and between DoD and IC elements when feasible, and then to use National Security Agency (NSA)-approved—e.g., Type 1—encryptors, In-line Network Encryptors (INEs), or traditional bulk encryptors to protect classified data transported over networks. To protect sensitive data exchanged among unclassified Local Area Networks (LANs), the strategy is to use commercial solutions that: satisfy published criteria; are validated by an approved, independent laboratory; are properly configured; and are accredited for use by an approval process such as the Department of Defense Information Technology Security Certification and Accreditation Process (DITSCAP).

For voice networks, a number of strategies are in use. The DoD's protection strategy is to use approved common user networks when available, or NSA-approved subscriber voice terminals otherwise. The strategy for DoD tactical networks is to use NSA-approved tactical radios, tactical subscriber terminals, or INEs to protect classified information transmissions. Law enforcement organizations use encrypted communications in the field, generally following Na-

tional Institute of Standards & Technology (NIST) Federal Information Processing Standards (FIPS) publications on encryption standards. Other civil agencies involved in tactical operations, such as responding to natural disasters, generally use Commercial Off-The-Shelf (COTS) communications with no encryption. They are migrating to digital phones, which are less likely to be compromised. However, this move is motivated by market changes rather than any requirement to have more secure communications. The most critical requirements for emergency response functions are availability and reliability, not confidentiality.

To achieve interoperability between government and commercial networks, the strategy is to include denial of service protection measures in all service level agreements for commercial leased network services. For DoD owned and operated networks, the strategy is to provide a number of measures to ensure network availability. These measures include: mechanisms that ensure the positive control of network elements; Public Key Infrastructure (PKI)-enabled authentication and access control for remote management of all critical network elements; authentication and integrity protection for all network management transactions; and enclave boundary protection for centers that manage the control of DoD WANs.

The Defend the Network and Infrastructure chapter of the IATF consists of several sections. The availability of backbone networks section considers data communications networks, e.g., Internet Protocol (IP) and Asynchronous Transfer Mode (ATM); and issues with secure network management. The wireless section considers the security issues associated with cellular service, pagers, satellite systems, and wireless LANs. The system high interconnections and Virtual Private Networks (VPNs) section addresses secure connectivity between systems operating at the same sensitivity level via backbone networks. A future section dealing with secure voice transmission will cover voice over the PSTN, voice over Integrated Services Digital Network (ISDN), and voice over data networks. A future section on multiple security layers will address issues with using a single backbone to transmit information of the same classification level, but of varying compartments.

This page intentionally left blank

## 5.1 Availability of Backbone Networks

Reliance on commercial providers of network services has been increasing, primarily owing to increased competition after the Telecommunications Act of 1996 and the exponential demand for bandwidth. While most private sector organizations traditionally relied on commercial providers for almost all of their network services, Government took a different view. Many Government organizations, especially the DoD and IC, held to the paradigm that they had to operate and maintain the entire communication system, including all of the long haul communication transport systems.

With the move to more cost effective commercial service providers, Government organizations have had to join private sector organizations in seeking to influence the network security industry. The overall strategy for the public and private sector should be first, to educate—organizations should understand the different aspects of network security and determine their own requirements. Secondly, they should seek to participate in standards activities to influence standards, protocols, and operations.

This section of the framework focuses specifically on improving the availability<sup>1</sup> of the long-haul transport systems to meet the operational requirements even if the long haul transport systems are under an Information Warfare Attack.

### 5.1.1 Target Environment

This section of the framework focuses on BNs. The most common examples of a commercial are the terrestrial-based voice systems and the Internet. In the DoD, the most common data BN is the Defense Information Systems Network (DISN). The framework looks to encompass a wider range of systems, than data Wide Area Networks (WAN) (including wireless systems, satellite systems, video teleconferencing systems, and voice systems). *BNs* hereafter refer to this entire range of communication systems.

Typically, BNs are known by a single name, such as the Internet or the DISN. However, these networks are constructed of a range of technologies and transport systems. Although the separations between BNs and other parts of the communication systems are neither simple nor clean, nevertheless useful characteristics that can be described in terms of a generalized model of a BN. We can decompose our model of the BN into nine focus areas:

- Network-to-network communication.
- Device-to-device communication.
- Device management and maintenance.
- User data interface.
- Remote operator-to-Network Management Center (NMC) communication.

---

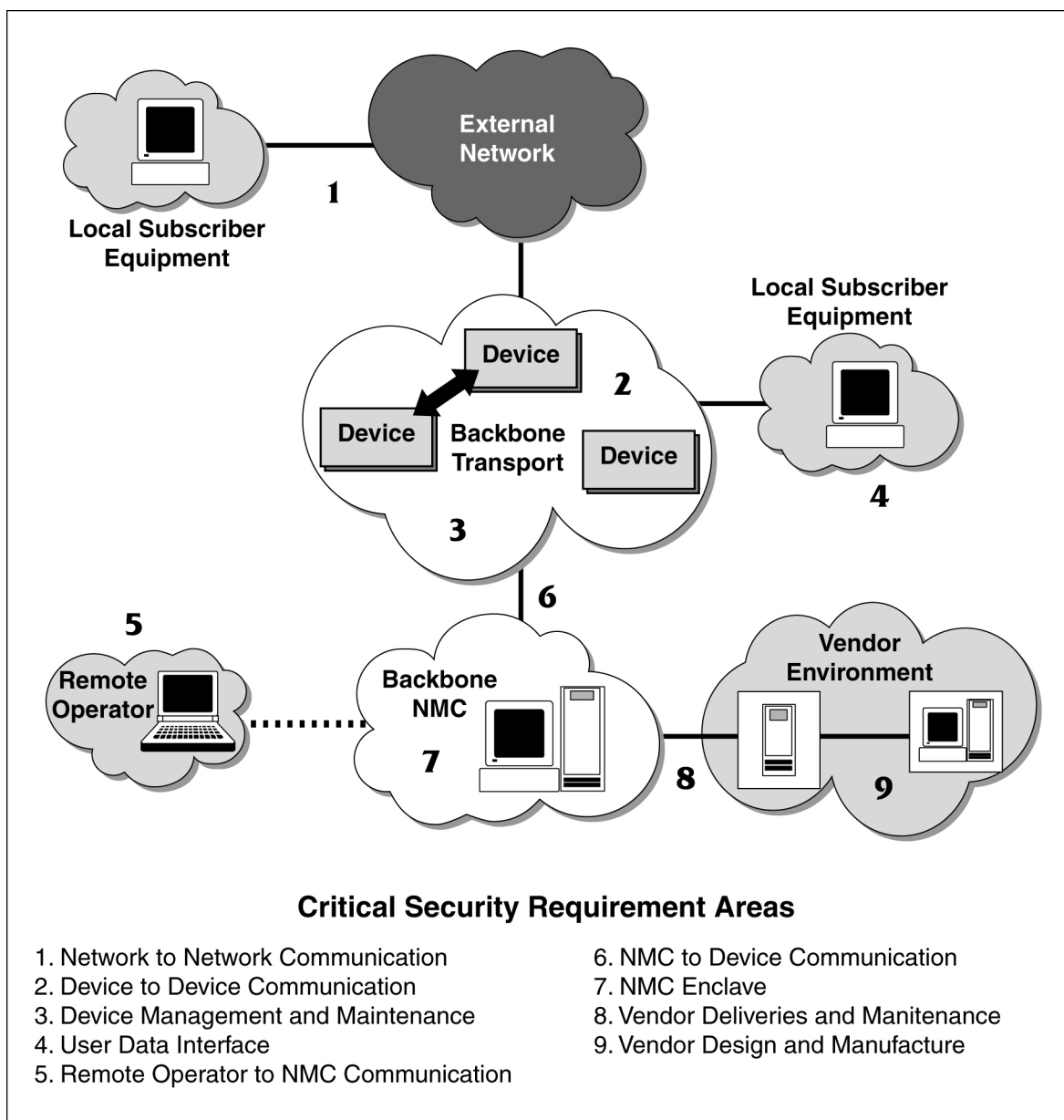
<sup>1</sup> The backbone security service is limited to availability for two reasons. First, backbones may be acquired through commercial service provisioning thus restricting the acquisition office from dictating special security services. Second, the communication models used in today's systems dictate the other security services, such as confidentiality and data integrity, to be handled by the end system and not the backbone network.

- NMC-to-device communication.
- NMC enclave.
- Vendor deliveries and maintenance.
- Vendor environment.

The availability of a BN is closely connected to the communications between networks, network devices such as routers and switches, and the network management's centers and the devices they manage. Additionally, the NMCs and network devices must be protected. We performed an Information Security (INFOSEC) Information System Security Engineering (ISSE) analysis over the model components for five network cases. The remainder of this section presents the backbone availability model and security issues.

The following provides an expanded description of the nine backbone availability model components identified in the model depicted in Figure 5.1-1.

- 1) *Network-to-Network Communication*—There are two classes of network traffic or data of concern here. One data class is the user traffic or user data that traverses this interface. The other data class, control traffic, is the communications required between the backbone transport devices and the external network devices. It is necessary to distinguish between two classes. Typically, the device-to-device communication is a well-defined protocol providing network-specific data necessary to transport the user data. The user data will be entering and exiting the backbone transport network. This is one of the BN boundary interfaces that allow the ISSE to define the inside and the outside of the BN.
- 2) *Device-to-Device Communication*—This area considers the internal communications between devices that are components of the BN itself. Generally, BNs require continual information exchange of this management and control traffic among devices to provide optimum performance and to support on-line maintenance and troubleshooting.
- 3) *Device Management and Maintenance*—This area focuses on configuration and parameter adjustments required to maintain the individual devices on the BN, the network management traffic. Typically, each device has a unique set of operational requirements and specifications that must be controlled by the NMC or maintenance personnel in order for that device to remain an active node on the network.
- 4) *User Data Interface*—The user data interface is the means by which user data enters and exits the BN. This may occur at any connection supporting user connectivity including user networks represented by the Local Subscriber Environment (LSE) and other networks connected to user networks represented by the external network. These interfaces should be resistant to cyber attacks from the user connections.
- 5) *Remote Operator-to-NMC Communication*—The primary concern with this area is the operator's physical security, e.g., where the equipment, usually a lap top computer, is being used and what protection is afforded to the equipment. In addition to those security concerns, there is the connection into the NMC and the type of security needed to protect it. When this area is needed to support operational requirements, it increases the complexity of analyzing the NMC; so perimeter security considerations regarding access to the NMC should be analyzed.

**Figure 5.1-1 Backbone Availability Model**

- 6) *NMC-to-Device Communication*—Addressing this area allows analysis of the perimeters of the Backbone Transport and the NMC, recognizing the NMC requires connectivity to the devices making up the Backbone Transport for all of the management operations. The connectivity may occur through in-band or out-of-band signaling using either primary or secondary channels. This provides opportunity to access the BN devices, and the NMC equipment and data, plus it exposes network management data.
- 7) *NMC Enclave*—The concern in this area stems from the concept that network management is critically important to the availability of the BN and should be operated separately from, what has been called in this section, user data. The management equipment and data require security protection from attack so they may successfully perform their

mission, which is to manage the BN. Considering this as a local network environment will permit the ISSEs to take full advantage of virtually every other section of this Framework document.

- 8) *Vendor Deliveries and Maintenance*—This area is more complex than Figure 5.1-1 depicts. The NMC may receive equipment or software to be prepared for installation in the Backbone Transport. It is possible that the vendor will be called upon to provide product service and maintenance directly to the backbone transport devices. The NMC may receive the information from the vendor either indirectly, e.g. by the postal system or directly on line through a network connection. The ability to ensure the validity of the information and equipment received plays an important role in the availability of the BN.
- 9) *Vendor Design and Manufacture*—This area covers the entire manufacturing process from development to production to delivery of the end item, whether it is a device or software. Security must be applied over all of this so that what “comes out of the box” can be trusted to operate properly. Security must also be designed into the product so that many of the security requirements raised in the other eight areas can be achieved.

Now that the BN focus areas have been described, it is useful to return and discuss its generalized use and operations. One of the general characteristics of BNs is that it has an inside and an outside. The user community generally connects from outside of the *backbone transport* portion of the BN. All internal connections are either between internal parts of the Backbone Transport or with the *backbone NMC*. By extension, the NMC is considered to be inside the BN. In today's environment of searching for cost reduction while improving user services, a BN will likely interoperate with one or more *External Networks* in addition to the user community it supports. The external networks are typically deemed untrustworthy with respect to the BN being analyzed.

Another characteristic of a BN is that it is viewed by users as a means to an end for their missions. The user's requirement is normally to communicate with another entity, not the BN itself. In other words, the user information travels across the backbone but does not stop there. In Figure 5.1-1, the users are represented by the LSE clouds. The security concerns of the LSE are addressed elsewhere in this framework, e.g., Chapter 6, Defending the Enclave Boundary/External Connections.

In this model, the backbone transport devices are managed and operated remotely by the NMC using Commercial-Off-The Shelf (COTS) or Government-Off-The-Shelf (GOTS) network management systems. The NMC devices are separate and distinct from the backbone transport devices. It should be noted that the NMC component of the BN architecture is fundamentally the same as an LSE. Though the purpose and function may be different, the NMC architecture takes advantage of the appropriate security guidance provided throughout the rest of this document.

Generally, the NMC must be operational 24 by 7 (24 hours a day, 7 days a week). Because of that need, NMCs may support remote operator connectivity, represented in Figure 5.1-1 by the Remote Operator. It is common practice to provide remote access to system experts so they do not have to be physically present at the NMC at all times. A Remote Operator is similar to a generic remote user and some of the security considerations are the same. Please refer to Section 6.2 (Remote Access). However, a Remote Operator has a significant difference. A remote user connects into the backbone network either from a special service provision—e.g.,



roaming user dial-up service—or from some external network or LSE connection. The remote user is considered to be **outside** the BN. In contrast, a Remote Operator—who connects into the backbone NMC via a similar manner—is considered to be **inside** the BN.

In the full life cycle of a BN, new capabilities and features are constantly being incorporated into the devices that comprise it. Occasionally new devices or components are installed to replace or upgrade the existing devices or to expand the network and its capabilities. The security concerns associated with this evolution are represented in Figure 5.1-1 by the vendor environment and interface. A common practice in the network industry is to develop the devices and the product software/firmware and then ship these new components to the field in the same manner used by any computer-based product. One method that is often used is to post the product software on an Internet web site for customer downloading. This distribution approach is open to compromise. To maximize the availability of the BN, it is necessary to have trust (in a security sense) in the entire life-cycle process of the BN and its components.

## 5.1.2 Consolidated Requirements

The fundamental requirement for availability of BNs is that they are required to be present and functioning properly when the missions require them. The President's Commission on Critical Infrastructure Protection acknowledges the importance of solving this problem with the following: “The critical infrastructures [including Information and Communications Industries] are central to our national defense and our economic power, and we must lay the foundations for their future security ... ” [1] Specific requirements are identified below.

### Functional Requirements

- BNs must provide an agreed level of responsiveness, continuity of service and resistance to accidental or intentional corruption of the communications service. (The agreement is between the owners of the network and the users of the network.)
- BNs are not required to provide security services of user data (such as confidentiality and integrity)—that is the user's responsibility.
- BNs must protect against the delay, misdelivery, or non-delivery of otherwise adequately protected information.
- BNs, as a part of the end-to-end information transfer system, must provide the service transparently to the user.
- As part of the transparency requirement, the BN must operate seamlessly with other backbones and local networks.

#### 5.1.2.1 Security Requirements

##### Access Control

- Access controls must be used to differentiate access to the network devices between users' access for transport of data and administrator access for network management

and control. For example, access controls must enforce user's access to status information versus configuration information.

- Access controls must limit access to the NMC.

## Authentication

- Network devices must authenticate the source of all communications from other network devices, such as routing messages.
- Network devices must authenticate all connection requests from network management personnel.
- Network management systems must authenticate network management personnel prior to being granted access.
- The NMC must authenticate the source of all communications entering the NMC from external networks.
- The NMC must authenticate the source of vendor supplied material.
- The NMC must authenticate the source of vendor supplied software. For example, new releases of operating systems must be authenticated prior to being implemented across the network.
- The NMC must authenticate all dial-in users prior to granting them access to the NMC.

## Availability

- Hardware and software resources (such as user agents and servers) must be available to users.
- The service provider must provide a high grade of system availability for users.

## Confidentiality

- The confidentiality of key material must be protected.
- The network management system shall provide confidentiality of routing information, signaling information, and network management traffic to provide traffic flow security.

## Integrity

- The integrity of communications between network devices must be protected.
- The integrity of the hardware and software in network devices must be protected.
- The integrity of communications between network devices and the NMC must be protected.
- The integrity of vendor-supplied hardware and software must be protected.
- The integrity of dial-in communications to the NMC must be protected.

## Non-Repudiation

- Network personnel must not be able to repudiate changes to the configuration of network devices.
- Vendors must not be able to repudiate vendor supplied or developed hardware or software.

### 5.1.2.2 Networking Environments

Please refer to Section 5.3 (System High Interconnections and Virtual Private Networks(VPN)) of the Framework, where these requirements have been addressed in detail.

### 5.1.2.3 Interoperability Requirements

BNs must be able to securely interoperate with other BNs and local subscriber environments. This requirement includes the secure exchange of network management information and routing data.

## 5.1.3 Potential Attacks and Potential Countermeasures

As with the Requirements for Network Environments section above, please refer to the corresponding System High Interconnections and Virtual Private Networks (VPN), Potential Attacks, Section 5.3, for substantial, related material. The reader should note that this section has a somewhat different focus from that of Section 5.3. This section is focused on attacks against network management operations and against BN infrastructure devices. In addition this section focuses specifically on user data and information in terms of availability and delivery service capability in the presence of the attacks discussed below.

Threats to network availability can be grouped into three general threat categories as discussed below.

- *Loss of Available Bandwidth*—The threat category occurs because every network has a limited amount of network bandwidth. Attacks can reduce the amount of available bandwidth, limit network resources for legitimate users, and decrease the network's availability. These attacks generally do not impact the operational control of the network. The network is operating as designed and the NMC retains control over the network infrastructure. This category applies to model components 1, 2, 4, and 6 in Figure 5.1-1.
- *Disruption of Network Management Communications*—This threat category impacts the normal operation of the network. Intrinsically, every network must move information from one user to another over network communication channels. Attacks in this category threaten the normal flow of information through the network by disrupting the communication channels. Examples include shutting down circuits or providing erroneous routing information. These attacks focus on the network management traffic used to control the flow of information across the network. The network is not operat-

ing as expected due to the misdirection of the flow of information, but the NMC still has some control over the infrastructure. This category applies to model components 1, 2, and 6 in Figure 5.1-1.

- *Loss of Network Infrastructure Control*—This threat category is the most severe. These attacks represent a loss of control over the network infrastructure. Once the network managers have lost control over the network infrastructure, or over the NMC, they are no longer able to provide the intended network services and, in fact, the network assets are conceivably at risk of being used to support the attacker's goals. This category applies to Critical Security Requirement Areas (CSRA) 3, 7, and 9 in Figure 5.1-1, in terms of loss of control of the BN. The attacks may also occur via any of the other model components in Figure 5.1-1.

Each threat category represents a potential loss of network availability. However, the severity of the attack is related to the loss of control of the network, since control represents the ability of the network managers to respond to an attack. These categories are then considered within the context of the major threat categories discussed in Chapter 4 (Technical Principles) of the Framework.

The remainder of this section discusses the relationship of these three general threat categories and the Classes of Attacks described in Section 4.2 (Adversaries, Threats, (Motivations/Capabilities), and Attacks). Where appropriate, countermeasures for specific attacks are highlighted below. The countermeasures are consolidated in the section that follows.

### 5.1.3.1 Passive Attacks

Passive attacks monitor and collect information as they traverse the network. Previously, BN providers did not consider the passive intercept of network management data as a threat to the network except as a means of gathering information for a more serious active attack. An example was intercepting fixed Identifications (IDs) and passwords to support a subsequent attack on the control of the network infrastructure. Now, BN providers are viewing passive attacks with growing concern. Providers are considering the overall network topology as sensitive information, with its protection from passive attacks needed to mitigate potential disruption of network management communications.

It remains to be seen which way the commands, status and the rest of the network infrastructure management traffic will be viewed in the future, but it seems BN providers are working hard to improve security. This is demonstrated prominently in the latest release of the Simple Network Management Protocol version 3 (SNMP v3) which has significant security section additions over earlier versions of SNMP. This class applies to model components 1, 2, 4, 5, and 6 in Figure 5.1-1.

### 5.1.3.2 Active Attacks

Active attacks represent the classic attack by an outsider <sup>2</sup> on the network. In the case of the backbone availability model, the outsider is represented by a “user of the network” or by an

---

<sup>2</sup> Note that for the Availability of Backbone Networks section of the Framework views insiders and outsiders from the view of backbone networks. Thus, insiders are those authorized to control and manage the network; outsiders

adversary connected through an external network connection (as opposed to the insider who is the manager or administrator of the network). All three general threat categories identified in above in Section 5.1.3 (Potential Attacks and Potential Countermeasures) can be realized; the following discusses the general threat categories relative to this class. These attacks apply to model components 1, 2, 4, 5, and 6 in Figure 5.1-1.

**Loss of Available Bandwidth Attacks** – Network bandwidth represents the network’s ability to transfer information. Loss of available bandwidth attacks (the first general attack category discussed above) consumes network bandwidth, preventing legitimate network users from exchanging information. Three common Available Bandwidth Attacks are the following.

- 1) Jamming attacks are usually the easiest to detect and possibly the hardest to counter for a network backbone. For example, in a jamming attack an adversary transmits noise in the electromagnetic spectrum of the network preventing the flow of information. Two examples are between a satellite and a ground station or between cells of a wireless network.

A variety of countermeasures—e.g., frequency hopping and redundancy via an alternative media such as terrestrial-based hard-wired systems—for these attacks have been developed for military applications. These countermeasures are usually not implemented in commercial backbone BNs due to cost and other constraints. These attacks apply to model components 1, 2, 4, and possibly 6 in Figure 5.1-1.

- 2) Flooding attacks consume network bandwidth by “burying” the network with processing communications in excess of network capability. Everyone is familiar with the problems with the phone system over holidays or during disasters where everybody tries to use the limited resource at the same time. Active cyber flooding attacks produce the same result using spurious communications traffic. This attack is difficult to counter since the network managers can rarely distinguish legitimate traffic from spurious traffic. The two most common countermeasures are to support a preemption capability, which allows specified users the right to specific bandwidth regardless of other demands on the system, or to limit the bandwidth available through any access point onto the network. This attack is typically applied at model components 1, 2, and 4 in Figure 5.1-1.
- 3) Theft of service attacks may be the subtlest of the Available Bandwidth Attacks. These attacks consume bandwidth, but they appear as normal operations. Attackers pose as legitimate users, establish a connection, and use the network to transfer their information. Most of the time, network managers don’t realize bandwidth is being stolen until valid users receive their bill and claims that they did not make specific calls.

A typical countermeasure is to require the users to authenticate themselves to the network before being granted access to network services. Another countermeasure relies on audit techniques. For example, the system could maintain a profile of users’ normal activities and trigger an alarm when the network detects abnormal activity. This attack applies to model components 1 and 4 in Figure 5.1-1. It is also possible at model components 2 and 6.

---

include both authorized users of the network (who do not have privileges to effect the control of the network) as well as potential adversaries that do not have authorized access.

**Disruption of Network Management Communications Attacks** – These are active attacks that disrupt network communications, intending to interfere with the flow of information across the network by attacking the control commands to the infrastructure devices. By way of contrast, bandwidth availability attacks do not impact the normal operation of the network. They consume bandwidth, limiting the availability of the network but not modifying the command and operation of the infrastructure devices. Network managers are still able to control the network, but the network is receiving misinformation causing a disruption in service. For example, Internet Protocol (IP) routing networks pass network topology data between the routers. This data allows the routers to move a user's information across the network. If this data is modified, the routers no longer deliver the user's information as expected, reducing the availability of the network.

Attacks in this category are specific to the BN and how it establishes and maintains the communication pathways to transfer a user's data. For example, voice networks rely on Signaling System 7 (SS7) to manage voice circuits. An attack on this network is to insert a message signaling "one of the user's hanging up the phone" resulting in the circuit being dropped. Asynchronous Transfer Mode (ATM) networks establish virtual circuits to transfer a user's data. An example of a disruption attack on an ATM network would be to transmit an Operations, Administration and Maintenance (OA&M) cell telling a switch to shutdown the virtual circuit. Analysis of this area of attack considers model components 1, 2, 4, 5, and 6 in Figure 5.1-1.

Two countermeasures are available to protect against disruption attacks. First, all network management traffic should originate within the network. This countermeasure requires the network edge devices to check all traffic entering the network to ensure that no network management traffic enters the network from the outside. This approach is referred to as establishing a security perimeter, an enclave boundary, on the system. Second, the integrity and the authenticity of network management traffic should be verified. For instance, a digital signature could be incorporated into the network management traffic. This mechanism could also be used to protect against a replay of valid network management traffic with the incorporation of time stamps or sequence numbers into the signed network management traffic.

**Loss of Network Infrastructure Control Attacks** – The most severe attacks are those against the network operators' control of the network infrastructure. Three ways of attacking control of the network infrastructure are the following.

- 1) *Network control attacks directed at the communications between the network operators and the network devices*—These attacks seek to isolate the network operators from the network devices. For example, network operators may access their network through a single connection point into the network. If this point is compromised the network operators cannot access the network.

The best countermeasure is to provide redundant access to the network, allowing the free flow of information from the network managers and their devices. This countermeasure has implications later in this discussion for another control attack.

- 2) *Network control attacks directed at network devices*—These attacks focus on getting access to, and thereby control of the device. For example, most network managers remotely manage their devices and use Telnet or other communications protocols to log into the device. Once the network operator has access, the device can be re-configured, includ-

ing changing the password used to log into the device. An adversary may choose several ways to gain this control. One example is for an adversary to actively attack the access control using password-sniffing programs. Two possible countermeasures for this attack are to strongly authenticate network management requests prior to granting them access to the device or to set-up a protected channel, such as an encrypted virtual private network, between the network operator management station and the device.

- 3) *Network control attacks directed at the NMC*—If the NMC is rendered inoperable, then the network operators are unable to access, let alone manage the network. Every communications path into the NMC serves as a potential attack path. Viruses are an example of these attacks. Viruses could destroy the contents of the memory of the network management devices. Several types of countermeasures are available to protect the NMCs against these attacks. Network guards or firewalls can be used to monitor the communications entering the NMC. These devices can prevent unauthorized communications and check incoming traffic for viruses and other threats to the NMC. A second type of countermeasures is procedural. Policies and procedures should be implemented to support the restoration of the NMC or establishment of redundant NMCs.

### 5.1.3.3 Insider Attacks

The insider threat considers an insider to be any user or network management operator of the system who knowingly or unknowingly causes the reduction of the availability of the BN. Insider attacks are initiated by personnel responsible for managing the network. The majority of these personnel are located in the NMC. In the analysis of BNs there are two “insiders.” There are the operators of the network represented in the model by the backbone NMC. The model recognizes a special case of management personnel: the personnel that operate remotely from the NMC and require additional scrutiny. There are also the developers and producers of the network components, represented in the model by the vendor environment. Specific insider attacks relevant to BN availability include the following.

- Backbone NMC insider has direct access to the NMC management assets. These users have legitimate reasons for accessing and configuring network assets. These users have the ability to launch subtle attacks on the network, by supplying misinformation to the network assets, or blatant attacks by transferring control of the network assets to an outsider.
- The most effective countermeasures rely on strong procedural mechanisms and strong accountability. Procedural mechanisms can be implemented to separate critical network functions, such as the configuration, maintenance, and provisioning of network assets from non-critical functions such as general e-mail and web-surfing. Audit mechanisms can be implemented to review the execution of network operations.
- *Remote operators* are a special case of the backbone NMC insider. These operators are generally on-call experts who help troubleshoot network problems. These operators pose as big a threat as the normal backbone insider does, but their identity cannot be confirmed by procedural mechanisms, and their commands can be compromised during transmission.
- A common countermeasure is to employ a secure modem to protect their dial-up connection. Regardless of the type of remote connection, the identity of the remote operator

should be authenticated and the integrity of the transmitted data protected. Analysis of this area of attack considers CSRA 5 in Figure 5.1-1.

- Vendors and producers that develop software control many, if not all of these devices. Commercial software is not typically developed with the strict configuration control that is associated with the development of trusted software. Therefore, there is a potential that malicious code can be embedded in the software. This code can support a range of attacks on the network infrastructure including the destruction of the system configuration information, the generation of spurious command information, and the loss of control of the network devices. This threat recognizes the malicious intent of the code inserted into the operating system; another aspect that must be considered is development software that could be exploited. Software developers are infamous for inserting “backdoors” and other features that allow to easy access to the system they are working on. If these undocumented features are not removed before the software is released, they could be exploited by an outsider to gain control of the system.
- The most effective countermeasures to this threat are procedural mechanisms. These mechanisms include the implementation of a strong software engineering process, which identifies the requirements for every software module and reviews the implementation, and strong configuration management. Analysis of this area of attack considers model component 9 in Figure 5.1-1.

### 5.1.3.4 Distribution Attacks

Distribution attacks alter the hardware and software provided by the vendors (commercial or government) as the mechanism to attack the network. These attacks are not limited to the vendor’s personnel, but include the delivery cycle as the hardware and software moves from the vendor to the NMC. The distribution threat needs to consider the movement of new software releases from the vendor to the installation in the network backbone. A common distribution mechanism is to provide a web server that users access to download the new releases. Currently, users cannot distinguish legitimate material from modified material.

An effective countermeasure is to apply digital signatures to the material allowing the network managers to verify the integrity and authenticity of the information. Analysis of this area of attack considers model component 8 in Figure 5.1-1.

## 5.1.4 Technology Assessment

BNs are not limited to a single technology. Typically, a BN is constructed using a variety of technologies. For instance, the DISN uses IP routers to connect subscribers to the BN. Connectivity between routers is provided by commercial leased lines, satellite links or ATM switches. This section assesses each of the common technologies used to construct a BN and addresses the available security features.

The technology assessment cannot be limited to the routers and switches used to pass data across the network; it also needs to look at the technologies used to manage the networks. In some instances, a single technology or technique can be used for a number of different types of devices, such as SNMP or Telnet. Alternatively, a single or proprietary protocol may be used to



manage the network devices. This section looks at the security features in network management protocols for Data Networks-IP Router Networks. Later releases of the Framework will look at the security features of Multimedia networks and ATM networks.

### 5.1.4.1 Data Networks IP Router Networks

IP networks are prevalent in today's commercial and government environments. IP network devices used in the wide-area infrastructure must have security features which promote a more robust and secure environment. IP is a connectionless packet oriented protocol that requires security considerations that are different than other technologies used for WANs. IP is a shared media so information that is addressed to a particular destination is readable by multiple network elements. Connections between peers may traverse multiple nodes or hops in the network. For security, this means that a network element does not know its immediate neighbors. Security services, i.e., authentication, access control; must be performed on a per packet basis, because a packet received on a port of an IP router may have originated almost anywhere in the network. Additionally, because IP packets are variable in length, security relevant information may be included with each IP packet.

## IP Transactions

There is network control and management traffic within wide-area IP networks that is required for the BN to function properly. Through the manipulation of these communications, an attacker may modify the operation of the network to accomplish his goals. Because IP is a very dynamic environment, packets may be misdirected, directed through specific routers, or service may be selectively or globally denied. The following sections describe the IP network communications that require security enhancements and which security services can provide protection.

## Domain Name Server

IP networks are dependent upon translating high-level domain names to IP addresses. This service is dependent upon the information stored on local and regional Domain Name Servers (DNS) to be accurate. Without accurate translation between domain names and IP addresses, IP packets cannot be properly routed through the network. Connections will either not be established, or established to end systems other than the intended end systems. The DNS query contains address information that must be translated as well as the responses to previous translation requests.

The integrity of this transaction is essential to establishing communications with the intended end system. The information on the DNS server, as well as the DNS query must not be modified by an unauthorized operator. One of the basic design philosophies of DNS is that DNS information is public and should be provided to all inquirers. Therefore there should be no attempt to implement an access control policy for DNS. Authentication and integrity are critical for an inquirer to know that they have contacted an authorized DNS server, and that the information retrieved from the DNS server is accurate.

## Internet Control Message Protocol

To report errors and unexpected error conditions, or to support network functionality, Internet Control Message Protocol (ICMP) is included with all IP implementations. ICMP poses several unique problems. ICMP messages may be viewed by any node within the network, and it is local policy for each node to act or not act on an ICMP message that it has seen. Additionally, ICMP is an IP layer protocol and does not ride on top of Transmission Control Protocol (TCP) or User Datagram Protocol (UDP). ICMP messages terminate directly at the operating system kernel and are not passed up the protocol stack.

ICMP messages should not be encrypted because all nodes in the network must be able to view them. ICMP messages must only be acted upon when they are received from an authenticated source. Additionally, ICMP messages must also pass an integrity check, to verify that they have arrived as intended. However, there are no security solutions implemented or under development to solve the problem of unauthorized ICMP messages. The recommended approach for local enclaves is to filter on ICMP messages and to only allow those ICMP messages that are critical to operations. This approach does not eliminate the risk of ICMP unauthorized ICMP messages, but it does reduce the risk. In WANs this approach is not viable. The WAN may need to transport ICMP messages between enclaves. To meet customer requirements for supporting network services, filtering on ICMP messages is not an option.

## Routing Messages

An essential part of any IP network, is a dynamic routing mechanism to efficiently transfer packets through out the network. The accuracy of these routing messages as well as the routing tables stored on routers is essential. This accuracy ensures that the routes that the connections take through the network are not denied and make effective use of network resources. Protecting a router's routing table is critical to preserving the availability of the network.

Integrity mechanisms are required for the routing updates sent between routers. This will ensure that routing updates are not modified as they travel through the network. Internal to the routers, an integrity mechanism is also required. Routing tables must be protected against unauthorized modification to ensure that they contain an accurate representation of the network. Additionally, an authentication mechanism is required to ensure that routing updates are not being injected into the network from an unauthorized source.

## Boot Protocol/Dynamic Host Control Protocol

The Boot Protocol (BOOTP) protocol is used when a network device powers up and needs to determine its IP address and possibly its hardware address. If a BOOTP message is intercepted en route to the BOOTP server, an attacker may respond with their own reply. This may cause the network device to download the incorrect memory image, which could have improper configuration information. The Dynamic Host Control Protocol (DHCP) extends this capability to allow dynamic IP addressing. Addresses of other necessary network elements, i.e., location of DNS server, location of timeserver; may be contained in a reply to a DHCP request.

The security services required to protect BOOTP and DHCP messages are authentication and integrity. Integrity ensures that BOOTP and DHCP replies are not modified while traversing the network. It is also important for the BOOTP/DHCP server to authenticate itself to the net-

work device to ensure that an attacker is not masquerading as the BOOTP/DHCP server. Configuration information received in a BOOTP/DHCP response must be received from an authorized server.

## Network Management

Perhaps the most critical area for WAN availability is network management. IP devices must be configured properly and must be resistant to malicious or unintentional tampering in order to provide network services. There are several physically different methods of managing an IP device. These are:

- *Inband*: Network manager connects to the network device using the same communication channels used for user traffic. The protocols used for this may be SNMP, Telnet, or HyperText Transfer Protocol (HTTP) for web based management.
- *Ethernet Port*: Network managers connect to the network device using an Ethernet network physically separated from the network used for user traffic. This requires an additional network infrastructure to support management traffic. The protocol used for this may be Telnet, or HTTP for web based management.
- *Local Port*: Network managers connect to the network device via a local port, i.e., RS-232 port, on the device using a laptop or similar computer. This method usually requires the network manager to be in close proximity to the network device. The protocol used for this may be Telnet, or HTTP for web based management.
- *Modem Port*: Network managers connect to the network device remotely using a modem interface on the device. Communications are usually over the Public Switched Telephone Network (PSTN) and operators may dial in from remote locations. The protocol used for this may be Telnet, or HTTP for web based management.

There are several security services that apply to secure network management. The first line of defense for network management is authentication. Administrators must first authenticate themselves to the network device to prove they are who they claim to be. Closely coupled to authentication is access control. Once an administrator's identity has been proven, their privileges must be determined. There should be several administrative roles on each device, each role with its own set of privileges. This allows each administrator to perform their job duties, but does not grant global privileges to each administrator. An audit log that links administrators to events and the time those events were performed is important. Such an audit log provides a mechanism for determining if a security violation has occurred, who is responsible, and suggests precautions for preventing similar events in the future. Finally integrity is important to ensure that communications between network managers and network devices are not altered while traversing the network. It is critical that configuration files on the devices are not modified by unauthorized personnel.

Traffic flow security for network management traffic may be of concern to some organizations. Network management traffic contains addresses of network components, or other information that may be sensitive. Providing confidentiality for network management traffic will provide protection for information while in transit through the network.

## 5.1.5 Framework Guidance

Our analysis of BN availability has resulted in some general guidance. This guidance is applicable to all of the network technologies that should be implemented to protect the availability of these networks:

- *Protection of Network Management Communications*—While the content of network management traffic is not considered critical, the integrity and authenticity is critical. Digital signatures or some form of secure hashes should be incorporated into all critical network management traffic. These communications also include the vendor-supplied software used to manage the network assets. If traffic flow security or disclosure of information within the network management traffic is a concern, confidentiality should be provided.
- *Separation of Network Management Data*—Backbone availability is not dependent on the protection of user data, but it is dependent on the protection of network management traffic. Countermeasures should be employed to isolate network management traffic from user data. One mechanism is to use an out-of band or dedicated communication channel, such as SS7. The value of separating management traffic from user traffic is to allow the infrastructure to provide the appropriate protection to the user data while impacting network performance only minimally. Network management data should be separated from the user data, and should be protected cryptographically. There are several means available for providing this protection, including encryption, digital signing, and cryptographic checksums.
- *Protection of the NMC*—The NMC is the critical element for maintaining control of the network. As long as the NMC can access the network, the network managers can respond to attacks. The NMC should be protected using the appropriate procedural and physical controls, and network security devices. A security device commonly employed today is a firewall. The NMC should consider constraining its operations to the management of the network. Permitting duties or capabilities beyond that which is necessary to manage the network provides a potential point of attack against the NMC.
- *Configuration Management*—System owners and operators should adopt formal configuration management practices. Strong configuration management allows network managers to restore network operations quickly and effectively after an attack. Configuration management supports the proper implementation of new releases of network software and the implementation of security upgrades. Strong configuration management also protects new releases of network software as the vendors develop them. Finally, it supports rigorous security design and analysis of the system.

The following section provides guidance for the protection of IP data networks. As technology assessments are completed for the other data networks, matching guidance will be incorporated into the Framework.

## 5.1.5.1 IP Data Network Guidance

### Routing Security

There are commercial implementations of cryptographic checksums applied across routing update messages.

### Address Space

Some government sponsored WANs may have the requirement to protect the addresses of the network elements. To accomplish this static routes must be configured between the WAN and each adjoining network. Network Address Translation (NAT) must be configured at the wide area border node to hide the addressing scheme of the WAN. Conversely, the local network may have the requirement to hide their address from the WAN. In this case NAT must also be configured at the local border node.

In the case of a public carrier network as the WAN, the addressing scheme may not be able to be protected.

### Filtering

Filtering, as it is traditionally thought of, is generally not applicable to WANs. Services cannot be filtered because it is likely that every service will be required by at least one user network. However, filtering is applicable to the area of network management. Each network device should contain a list of identifiers that describe the administrators with configuration/viewing privileges on that device. This has historically been done on IP address. IP addresses are easily spoofable. Another mechanism in addition to IP addresses is required to determine which administrators are capable of modifying/configuring each device.

### IP Security

IP Security (IPSec), as defined in RFC 1825, is a set of protocols supporting the secure exchange of packets at the IP layer. To achieve this, IPSEC employs two cryptographic security mechanisms: the Authentication Header (AH) and the Encapsulating Security Payload (ESP). These IP-layer security mechanisms may be used together or separately. IPSec is currently being incorporated into vendor products. IPSec functionality should be available in commercial IP network elements.

While IPSec is a suitable set of protocols for providing confidentiality for user traffic, it was not designed to provide security for intra-network communications. IPSec may be used to implement some VPN scenarios required for segregation of user traffic over the WAN. IPSec is not viewed as being able to provide security to achieve WAN availability.

### Network Management

**Inband**—Inband network management is performed using SNMPv1. There are no security features inherent to SNMPv1. All Management Information Base (MIB) information must be considered accessible to an SNMP agent. Devices typically employ IP address filtering to limit

the management stations that may configure/manage the devices. While it is recommended that this feature be used in WANs, it is not sufficient to prevent unauthorized access to network resources. IP address spoofing is common and easily implementable. The recommended approach to inband network management is SNMPv3. SNMPv3 provides confidentiality, integrity, and authentication, and timeliness functionality to inband management.

**Ethernet Port**—Constructing a separate Ethernet network to provide network management is a secure method of network management. It is a physically separate network, which provides a larger degree of control of the network management network. However, for wide area networks, this approach is not practical. The network elements are geographically disperse and it not feasible to construct another WAN for management. If Ethernet port management is not being used, it is recommended that the network device be configured to disallow network management connections through the Ethernet port.

**Local Port**—It is critical that IP network elements can be securely accessed through a local port. This is often the network's configuration method if the BN element cannot be reached through the network. Physical security of the devices is important to protect the local port. If an attacker does not have physical access to the device they cannot be successful. Authentication and access controls are also critical. There should be several different administrative roles on the network elements. When administrators authenticate themselves to a device, they must assume a role with well-defined privileges.

## References

---

1. “The President's Commission on Critical Infrastructure Protection - Report Summary”,  
[http://www.info-sec.com/pccip/pccip2/report\\_index.html](http://www.info-sec.com/pccip/pccip2/report_index.html)

This page intentionally left blank



## 5.2 Wireless Networks Security Framework

The Wireless Networks Security Framework section has been added as an element of the Information Assurance Technical Framework (IATF) to discuss the security of new “wireless” communications technologies. This section is incorporated because the IATF addresses many security concerns and secure infrastructure elements that also affect wireless communications. Exposure of wireless communications in the Radio Frequency (RF) transmission environment, and the portability of computer processing and storage that wireless connectivity provides, add another set of vulnerabilities to the vulnerabilities of wired network systems. This section will present the areas of security where wireless communication presents additional vulnerabilities, different customer requirements, and different, although related, security concerns.

Wireless network protection addresses the need to ensure security of user communications where one or more links in the communications channel traverse a wireless link. “Wireless” is defined as the set of services and technologies that does not include more traditional legacy radio communications such as Land Mobile Radio (LMR) and military point-to-point and netted Military Satellite Communications (MILSATCOM). RF systems are addressed separately because the Government legacy systems were typically designed for specific applications and included required security mechanisms. The new wireless technologies are commercially based, and are not built to specifications for Government applications, although the Government applications for such systems are increasing rapidly. Security measures for new wireless systems must be developed in conjunction with the equipment manufacturers and service providers involved in the wireless industry.

“Wireless,” in this context, defines a set of commercially developed systems, products, and a system infrastructure that transfers personal communications from wired to RF transmission environments. Wireless communications often are provided as a service to the user where the user does not own the communication infrastructure. These systems often do not require user licensing or user spectrum management (at least in the United States (U.S.)). Typically, wireless systems use low-power transmission and/or spectrum-spreading techniques in short range communications environments. The set of characteristics used herein to define wireless are:

- RF communications in commercial and unlicensed frequency bands,
- Low-power, short-range communications systems using enhanced processing and multiple transmitters to achieve range when required,
- Commercially owned and operated communication infrastructure (there are exceptions),
- Commercial standards,
- Vendor proprietary protocols, and
- Mobility of users and communications.

As we describe the technologies and applications involved in wireless systems, the reader will note that there are exceptions to each of these characteristics. Wireless communications, rather than being a set of discrete technologies, applications, and implementations, actually forms a continuum of capabilities that connect across the boundaries of the systems definitions we pro-

vide. Wireless technologies also, in most cases, rely heavily on the wired network and telecommunications infrastructures for their interfaces and proper function. These interconnections are significant in relation to discussion of security.

Wireless equipment may be used by traveler or telecommuters users to remotely access their Local Area Networks (LANs), enclaves, or enterprise-computing environments. However, most remote access situations involve connecting through wired telephone or commercial data networks. Discussion in this section of the Framework focuses on wireless communication networks in general, regardless of the systems being accessed through the network. As digital wireless telephony, two-way paging, Wireless LANs (WLANs), and other wireless technologies gain strength in the marketplace, both government and industry users are becoming increasingly more reliant on wireless communications for their daily activities. With this in mind, these devices must operate in untrusted, highly mobile, and potentially hostile environments.

There will be some overlap between the options presented here and in other portions of the IATF because the majority of wireless communications networks in use today tie into a larger, wired network with additional security concerns. Previous sections of the IATF have addressed the data network portion of these wired concerns in great detail, and references are made throughout this chapter to those IATF sections as applicable. Securing wireless communications across network segments implies a unique set of challenges that must be addressed within this Framework document in order to provide layered security services as outlined in the defense-in-depth strategy.

In today's marketplace, the consumer has access to a wide variety of wireless devices, including digital wireless phones, mobile satellite circuit-switched and packet services, WLANs, pagers, and wireless Private Branch eXchange (PBX)/local loop devices. Each device interacts differently with existing wired networks, often through a private gateway or service provider's network equipment. Additionally, different users have different connectivity and communications security needs. Information protection mechanisms can provide authentication and confidentiality, but definitely add to the cost of the equipment. Therefore, before purchasing any wireless communications equipment, users should make a decision regarding connectivity needs and the sensitivity of the information that will traverse their wireless network. Based on these decisions, appropriate protection mechanisms can be selected to meet user needs.

This section examines several categories of wireless technology, addressing the functional requirements and security requirements and mechanisms involved at each point in the communications and processing links. Security requirements will focus primarily on the following areas: Identification and Authentication (I&A), access control, data confidentiality, data integrity, and availability. These requirements for wireless systems do not replace those discussed in earlier sections. Instead, they are the same as the security requirements presented for wired networks, but may have differing emphasis due to RF exposure, and differing implementation requirements. For example, if a Sensitive But Unclassified (SBU) WLAN is connected to a public network such as the Internet, the requirements discussed in sections 5.3, 6.1, 6.2, and 6.3 are fully valid. RF transmission of sensitive or classified data adds yet another variable to the equation in terms of ensuring receipt of the message by only the intended recipient, detecting location of users, and denial-of-service through techniques such as jamming. In situations such as, a wireless network connection will often expand Virtual Private Networks (VPN), Protection of Network Access (PNA), Remote Access, and even Multi-Level Security (MLS). Typi-

cally, wireless systems connect to their wired counterparts at the same security level as the wired system, although the use of end-to-end confidentiality can permit users to “tunnel” through the wired system at any level of classification without mixing different classification levels. The provision of security mechanisms for high-to-low, low-to-high, and need-to-know are entrusted to processors within the system just as they are with wired components.

In developing the security solutions framework for wireless communications, we have subdivided commercial wireless communications into topical areas based on differences in application and implementation technology. Admittedly, there is overlap as providers merge applications to provide new services and maximize customer base, e.g., paging over cellular phones in Personal Communications System (PCS) networks. The wireless topics covered here are divided into the following areas.

- Cellular telephone.
- Low/Medium Earth Orbit (LEO/MEO) satellite telephone networks.
- WLAN.
- Paging (one-way and two-way).
- Wireless telephone (wireless PBX, Wireless Local Loop (WLL), and cordless telephone).

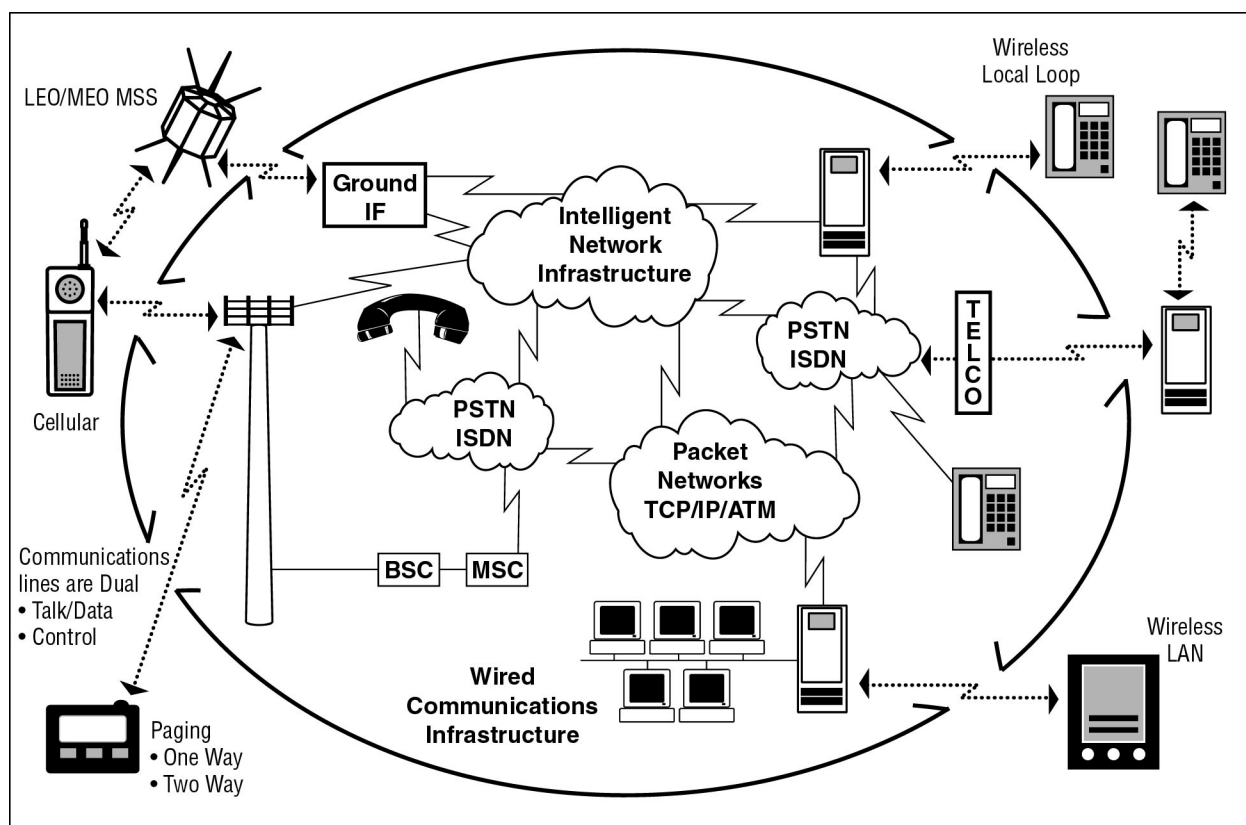
Figure 5.2-1 shows a combination of the wireless services attached to a set of wired infrastructures. It depicts a boundary around the various wired information transfer services that include both data network systems and circuit switched systems, which typically provide voice communications. Each type of wireless implementation effectively creates a hole in the wired infrastructure boundary because it exposes information in the system to the RF medium where signals can be much more readily detected and intercepted than in wired communications systems.

Figure 5.2-1 demonstrates that security measures implemented in the wired infrastructure can be negated by wireless connections. For example, a user community might have a wired Virtual Private Network (VPN) that is secured using a combination of encryption, access controls, and firewalls to create a security boundary shown as the oval in the figure. The connection of wireless components to the VPN (e.g., wireless LAN, cell phones) can expose the VPN users and their data to over-the-air signal intercept. Such interception is readily accomplished. The wireless assets, if not properly implemented, thus punch holes in the security boundary. These holes are depicted as the breaks in the oval in the figure.

Wireless technology and capabilities are moving so rapidly that continuous updates to this document will be required to attempt to stay abreast of increased bandwidths, new modes of wireless operations, new product and service offerings, and the aggregation of services. As wireless technologies services are enhanced, new vulnerabilities and user risks will be introduced.

Throughout this section, comparisons are made between several different types of wireless networks and their wired counterparts. New threats and new vulnerabilities in the wireless arena add a different dimension in security requirements and considerations for designers and consumers. Some of the vulnerabilities and risks described in this section of the IATF are common to both wired and wireless networks and communications media. This section will emphasize areas of risk that are increased by the use of wireless communications media. This

Framework will highlight critical gaps in current Government and commercial security technologies.



**Figure 5.2-1 Wireless Extension of the Wired Infrastructure**

## 5.2.1 Cellular Telephone

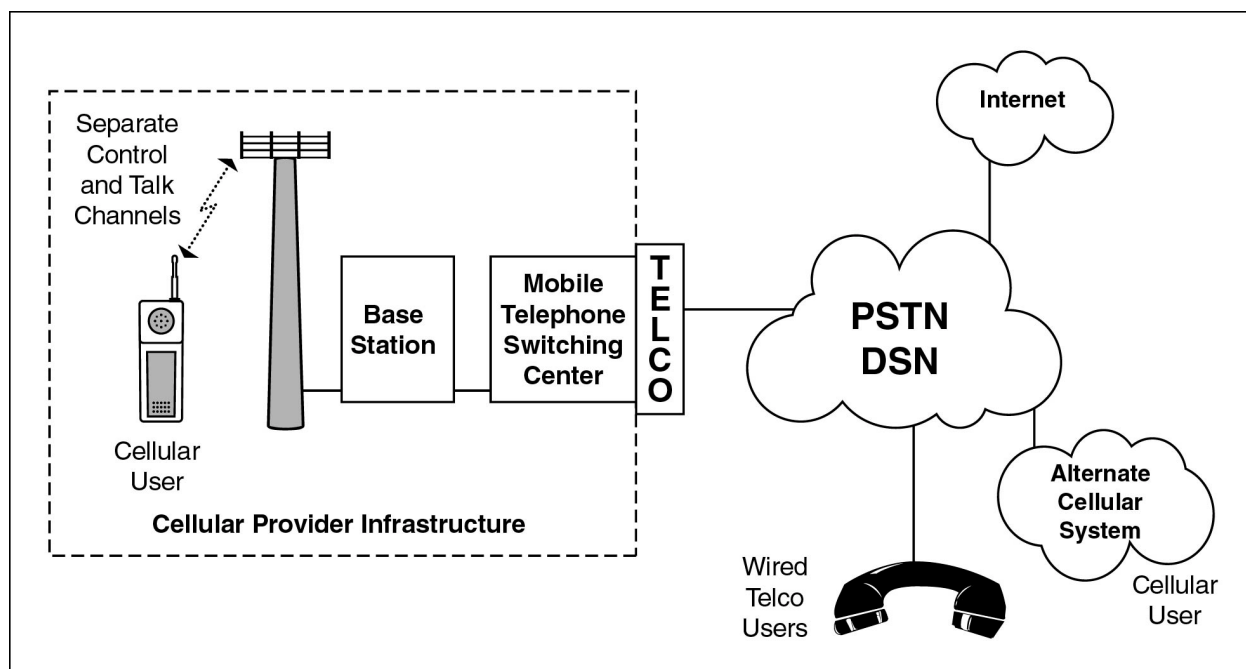
As technologies have advanced, cellular applications and terminology have become confusing as a result. Originally, “cellular” referred to a dialed analog voice telephone call technology that made use of distributed transceivers in line-of-sight communications with connections to the circuit-switched wired infrastructure. The term cellular no longer means the same thing for everybody because it is evolving into a digital pipeline that can be used for virtually any voice or data-based service (bandwidth limitations notwithstanding). Cellular systems operate primarily in the 800-900 MHz range and the 1.8-1.9 GHz range using either Time Division Multiple Access (TDMA) narrowband or Code Division Multiple Access (CDMA) wideband RF modulation. These distinctions of frequency and modulation do not substantially modify the services offered by cellular providers, but are in some cases germane to the security of the systems. All cellular systems provide an over-the-air control channel from the cellular base station in addition to multiple user “talk” channels. This arrangement means that the bulk of the system control is out of band with reference to user channels.

In recent years, the cellular telephone market has seen tremendous growth around the world. With the transition to digital cellular telephony and the advent of the new PCS, the wireless

telephone system has become a major part of both the Defense Information Infrastructure (DII) and the National Information Infrastructure (NII) for mobile users. Moreover, users desire similar functionality with wireless telephones as they have become accustomed to with standard wired telephones, including call forwarding, conference calling, and secure calling. Specialized militarized systems have been developed where vehicle transportable cell base stations are used as cellular telephone communications hubs. The user instruments that support cellular communications have grown increasingly capable in mobility, processing, storage, and communications capability. This aggregation of capabilities provides enhanced user functions, but also increases the risk of loss of sensitive information, denial-of-service, and spoofing of user messages and identities.

### 5.2.1.1 Target Environment

This Framework examines the standard wireless telephone environment, described as an end-user with a hand-held telephone, roaming throughout a cell-based infrastructure owned or at least controlled by a cellular service provider. As shown in Figure 5.2-2, the cell towers connect through a base station to their Mobile Telephone Switching Center (MTSC), which provides connection to the Public Switched Telephone Network (PSTN), or if service is procured, to the Defense Switched Network (DSN).



**Figure 5.2-2 Cellular Telephone Environment**

Figure 5.2-2 can be broken into three major sections: the user environment, the service provider network, and the public network. The user environment consists of the handheld phone and associated user, as well as the talk and control channels. The service provider network infrastructure includes all equipment and connections from the cellular transmitter through the base station, and on to the MTSC. The MTSC is the gateway to the public or defense-switched network for wired routing of calls. The PSTN includes connections to wired users, the Internet,

and to other mobile network providers. Each segment varies in the levels of privacy and availability provided to the user.

## 5.2.1.2 Consolidated Requirements

Users of wireless networks require functionality from their wireless equipment similar to what they get from their wired counterparts. Wireless telephony is certainly no exception. When discussing the following capabilities and postulated functional requirements, particular attention is paid to functions associated with connecting a wireless user to an existing wired network.

### 5.2.1.2.1 Functional Requirements

#### a) Users/User Equipment

- Should provide maximum portability and mobility for user.
- Must have individual Identification (ID); e.g., unique phone number.
- Must provide unique identification of user instrument.
- Must be able to provide location to the service provider system, e.g., Emergency 911 (E911).
- Must have service availability within full assigned area of provider network.
- Must ensure confidentiality of control channel and voice/data channel information.
- Must provide protection for information stored and processed in user equipment.
- Must provide user with maximum allowable access to needed information and services.
- Must be compatible with different signaling protocols for operation in different locations when outside home network.
- Must interface to wired and wireless user communities.
- Should provide certificate and key management and distribution interfaces for authentication of users.
- Should maximize user instrument operating time (battery life).

#### b) Geolocation is both a benefit provided by cellular systems (under certain circumstances) and a risk for cellular users when the function is not desired. Federal law for E911 service requires geolocation of users for emergency situations. At the same time the greater precision of the geolocation and the availability of that information in the cellular system puts other users at risk when clandestine operations are in process.

#### c) Service Provider

- Provide high grade of system availability for users.
- Provide high quality voice and error-free data services for users.

- Protect user information, e.g., ID and location, within the cellular infrastructure.
- Provide priority service for critical users.
- Provide capability for user communities to manage allocation of user services.
- Manage security of user provisioning and location information.
- Protect against the full range of network attacks, e.g., cloning, eavesdropping, impersonation.
- Provide signaling technologies that are compatible with multiple user instruments.
- Provide protection against jamming and other denial-of-service attacks.

d) Interface to Public Network

- Provide minimal operational impact to user and phone performance.
- Provide accurate billing method.
- Provide dedicated connections from Mobile Telephone Switch to Telco.
- Provide wired telco services; e.g., Caller ID.
- Provide standard interface to telco systems.

### 5.2.1.2.2 Networking Environments

- The networking environment in a wireless telephone network is not as clearly defined as it is in a computer network. One of the significant differences between a cellular network and a computer network is the level of access provided to a user. Local access to a computer network can provide universal access to all systems connected to that network. Access on a cellular network is much more limited for the end user, that is, access to a selected called party. However, with the increased use of the data capabilities of digital wireless telephony, a cellular network may begin to resemble the more familiar computer network. Wireless telephones should offer conference calling, as well as the ability to broadcast data to one or many recipients simultaneously.
- The networking environment should maximize the user's capability to use the service within the full boundaries of the service area. Fading and interference characteristics vary depending on site structures and modulation techniques. Users should investigate these characteristics for different providers in areas of critical operations for service continuity before selecting a provider.

### 5.2.1.2.3 Interoperability Requirements

- Service providers and associated handsets should not force users to use any non-standard protocols, modes of operation, or procedures that would prohibit interoperability with external users or systems to which users desire to communicate.
- Different cellular infrastructures currently make certain handsets inoperable in many areas around the world. In addition to the varying protocols, frequency allocations differ globally as well. While equipment is being manufactured to operate in different frequency bands, switching between protocols like TDMA and CDMA is more challeng-

ing. From a network security standpoint, users must carefully consider how transmitted signals affect detectability, availability, power control, jamming, and interception. Based on these considerations, the proper technology should be available to meet the user's needs. Regardless of the primary digital multiple access technique used, cellular handsets that can revert to a more universal system like Advanced Mobile Phone Service (AMPS) are extremely useful when the mobile user is outside of his normal area.

#### 5.2.1.2.4 Anticipated Future Requirements

- Convergence of technologies is demanding access to Internet services from the wireless telephone. Manufacturers have begun providing this service with a combination of wireless telephone and Personal Digital Assistants (PDAs). Increases in channel bandwidth to (in excess of) 100 Kbps within the next several years will make Internet connection a viable reality.
- Wireless phones will require operation with a smart card, or Subscriber Identity Module (SIM) card, for such future technologies as electronic commerce. These cards are also referred to as tokens. A token can be implemented in hardware or software, depending on the required assurance level for the transmitted information.
- Tokens will help cellular phones provide digital signatures, as well as end-to-end confidentiality of information. The security features required for electronic commerce can also be used to implement security features for sensitive and classified traffic.
- The ability to use a single-user instrument for different types of cellular protocols (and other wireless capabilities such as Mobile Satellite Service (MSS), paging, WLAN, cordless phone services, and wireless computer synchronization—all wireless services in a single device) is now coming on line. Universal handsets will be available in the near future. This will reduce the cost of confidentiality and other security mechanisms because the security will not need to be implemented for multiple protocols, but could rather become a user application that is independent of the network for end-to-end security requirements.
- The number of communications modes and interfaces described in the previous paragraph will require some common form of authentication and other common security solutions.
- Increased information transmitted over the user and control channels will require enhanced security for those connections. For example, Caller ID is now becoming available, and E 911 will carry very specific geolocation information over the RF path.

#### 5.2.1.3 Potential Attacks

The primary concerns of the cellular service provider are theft of service and denial of service. While different types of users may or may not be concerned about the confidentiality of the information transmitted and received by their wireless phone, commercial service providers definitely want to ensure the cellular system prevents unauthorized use of their service by a non-paying customer, and that the cellular service is functional for paying customers. Confi-



confidentiality of the information is typically a secondary objective for the service provider, but is a primary concern for business and Government users.

#### 5.2.1.3.1 Passive

- Eavesdropping operations were relatively simple with analog AMPS handsets. The change to digital technologies has increased the difficulty of passive, but devices can be readily modified to provide channel scanning and intercept capabilities. Without a true encryption scheme, passive is a major attack.
- Geolocation by an adversary via direction finding, cell location, or E911 requirements.
- Traffic analysis via dialed phone numbers and Caller ID.
- Spoofing. Attacker intercepts data, splices in information, and re-transmits the message as if originator of the message.

#### 5.2.1.3.2 Active

- As shown in Figure 5.2-2, a distinction must be made between the voice/information channel and the control channel. Intercept of the control channel information is a bigger threat to service providers, while users are typically more concerned with the confidentiality of the “talk” channel.
- Denial-of-service by jamming or altering control channel data can be a threat to users and service providers in cellular networks because of the vulnerability of control channel information when it is transmitted over the air. Such attacks typically require physical access to a provider’s network equipment, although outsider spoofing can modify the control channel.
- The threat of an outsider controlling the transmit power of the user handheld device. This capability allows an attacker to conduct locating and tracking operations against a target. Also, an attacker could cause a denial-of-service by limiting the output power of a user’s handset below what is required to maintain a connection.

#### 5.2.1.3.3 Insider

- Duplicate smart cards or SIMs (copy user token).
- Steal information on user identification and user traffic via control channel intercept.
- Modify control parameters of the system infrastructure.
- Modify user’s phone.

#### 5.2.1.3.4 Distribution

- Hardware or software modification in transit could be used as a first step in a complete attack by which an adversary eventually causes the system to send data or allow access by way of electronic connections to information for which he is not authorized. These attacks, however, are not the emphasis within this category.

- The distribution attack is enhanced by the fact that user instruments are becoming increasingly more modular. Thus, a user capability is assembled from parts that were distributed separately. Such components include storage devices (disks, flash prom), and communications devices, e.g., PC card modems, wireless modems, and WLAN cards that could spread viruses, and open undesirable communications channels.

### 5.2.1.3.5 Other

- Theft of portable user devices with sensitive information and user programs in storage. Because of the increasing integration of processing and communications elements in mobile systems, the theft of user equipment can be very destructive because of the storage volume and aggregation of information on that equipment.

## 5.2.1.4 Potential Countermeasures

Sufficient countermeasures must be implemented to provide privacy, authentication, and message integrity in accordance with the level of information being transmitted. Type 1 security, primarily for the DII community, requires countermeasures that provide the maximum possible security for message traffic. Sensitive information requiring Type 2/3 security requires the use of less stringent countermeasures. In order to maintain a secure infrastructure, the Government must overlay a supporting system infrastructure to incorporate authentication and key management and other countermeasures for each level of information as appropriate. Chapter 8 (Supporting Infrastructure) of this Framework is dedicated entirely to discussion of supporting secure infrastructure, and Section 8.1.5.14 (Attack and Countermeasures) covers attacks and countermeasures in more detail.

### 5.2.1.4.1 Encryption

The primary security requirement for cellular phones, as with any RF transmission system, is protection of user information over the air. There are two primary modes for protection. The first is encryption to secure the information, and transmission security, e.g., signal spreading or hopping; to protect the channel and possibly to provide protection against signal detection. Information on the control channel is also user-related at times in that it provides information on location, privileges, called party and calling party. Such information is very valuable for traffic analysis. A second important requirement for users is I&A of the parties in a communications session.

The Federal Bureau of Investigation (FBI) is presently promoting a law which will prohibit sale of encryption devices for use within the U.S. that do not provide key recovery services to support Communications Assistance for Law Enforcement Access (CALEA). Although not implemented, it appears that cellular service providers are slow to implement encryption services until the implications of a key recovery law are known. However, the techniques and standards for certificate and key management and encryption exist within the data network world to permit firmware or software encryption to be implemented for sensitive communications. Encryption algorithms can be embedded or implemented on the same tokens that provide user identification and privileges.

In-band signaling is also a target for encryption to prevent traffic analysis. For instance, encryption of dialing and data digit signals sent over the RF network must be considered, as well as Caller ID information that precedes a received communication. This will help secure credit card transactions, Personal Identification Numbers (PINs), other account numbers that are entered to access commercial dial-up services, and the identification of calling and called parties.

#### 5.2.1.4.2 I&A

SIM cards and other small token form factors may provide the best countermeasure to enable user and user terminal authentication (and security management). If a phone is stolen, for example, the user can notify the service provider, who then deactivates the SIM card in the stolen phone. The phone can even be programmed to flash “Stolen Handset” to notify the thief that the handset is useless. The same measures that providers use to prevent theft of service from the provider can be adapted to provide I&A security services. For increased security, service providers can permit user groups to control access of their own individuals using software tools that the service providers use to provision systems. The same provisioning capabilities can be expanded to include information such as security clearances, access to keying and other Security Management Infrastructure (SMI) services, and restriction of services within the limits of the overall provisioned (and paid for) service.

#### 5.2.1.4.3 Availability and Integrity

The availability and integrity of communications are largely a function of the protocols used by the service provider to connect calls, to provide reliable communications channels, and to service an optimal number of customers. As with any telephone system, busy channels are possible, although a system busy (rather than called party busy) is much more likely in cellular systems depending on the number of subscribers within a given cell or coverage area. To maximize the number of users in a given area, the RF power output is controlled for provider and/or user equipment on a dynamic basis to within a tolerable channel error rate for digital voice communications. Error correction codes are then used to correct the errors that would not be tolerable for data communications. To enhance both availability and integrity, a caller priority technique could be implemented to eliminate busy connections for critical calls and to reduce the number of concurrent general user calls processed within a given cell area in support of emergency operations.

#### 5.2.1.5 Technology Assessment

Within the wireless telephone market, current technology is more than adequate to permit insertion of required security to most applications, but few security measures have been implemented. As discussed earlier, the best available security technologies involve using some sort of token (physical component or inserted code) to provide authentication, access control, and data confidentiality. Lessons can be learned from the use of SIM cards with Global System for Mobile Communications (GSM) phones in the European market, where a user must have both a SIM card and a password in order to operate the telephone. Hardware or software tokens can be issued to every individual requiring sensitive communications who will use a wireless telephone in the future. Regardless of which protocol is used in a mobile telephone, the technology

is available to ensure that these tokens provide continued high performance and ease of use for the mobile user as well as providing a mechanism for implementing the required security. For U.S. Government applications, cellular end-to-end secure handsets are under development to satisfy Department of Defense (DoD) and other Government high-security requirements.

To manage the approval and provision of tokens and security privileges, a SMI infrastructure is required. Presently, the software cryptography implemented in some systems provides protection only for the lowest levels of assurance.

Communications bandwidths (typically less than 20 Kbps) are not yet sufficient to support efficient public key distribution capabilities over the cellular communications channels, but the picture is changing in two ways. High bandwidth cellular services (over 100 Kbps) will be coming on line within the next several years, and new techniques for key and certificate distribution based on elliptic curve cryptography will provide more efficient transfer mechanisms. In combination, these capabilities will minimize call set-up times and reduce the airtime cost of security to the point where a more widespread user base will consider its use.

### 5.2.1.6 Usage Cases

Other sections of this Framework have addressed several cases involving connecting equipment at one classification level to equipment at the same, or different classification level across both trusted and untrusted networks. These cases are clearly an IATF issue, but also apply in the wireless domain. However, use of wireless equipment interfacing with a wired network does not significantly change the cases that were previously discussed. In general, some level of communications security is recommended for any equipment where there is a connection to a potentially hostile or unknown environment. In the case of cellular communications, all transmissions can be thought of as connecting to an unknown environment because of the nature of RF transmissions and the ease of signal intercept. Thus, the descriptions of each of the specific cases addressed in this Framework remain unchanged for the wireless environment. Cellular calls are treated herein as having the same levels of classification as the wired systems to which they are connected. An exception involves the use of high-grade end-to-end confidentiality of the wireless service so that the user is independent of the classification level of the wireless or wired networks to which he/she is connected.

The cellular user scenario to be discussed is the voice phone call from/to a cellular portable phone system. Although the scenario appears to be quite simple, the actions required for the establishment and conduct of the call are quite complex. This “simple” example involves only a voice phone call; that is, no data, pager, or other service that might be available under services such as PCS.

There are three types of connections addressed in this scenario:

- 1) The cellular user calls a Plain Old Telephone Service (POTS) user.
- 2) The cellular user calls another cellular user (same or different provider).
- 3) The cellular user calls a satellite telephone user, e.g., Iridium phone.

The risks to users under the three scenarios are similar in terms of over-the-air exposure, but there are differences in denial-of-service and quality of service that must be considered. The

risks presented below will call out the specific situation under which a certain risk or degradation in service occurs.

It is important to note that any communications over commercial facilities opens up a large number of paths for the call control and user voice information to follow. The user has little to say about what path his/her information will take, or where important information related to the user will reside. As shown in Figure 5.2-2 for cellular voice calls, the paths that can be taken by a call are varied.

Before the user ever gets to the point of making a telephone call, the user has to establish service with a cellular provider. When the service is established, the parameters are set for local service areas and roaming areas as well as billing related items, e.g., free call minutes. All of these parameters are checked before calls can be completed. The user privileges can be checked rapidly by the provider through the use of the Wireless Intelligent Network (WIN) that provides a separate control system for the networks (separate from the cellular user channels themselves). User-related information is readily available within the cellular control infrastructure.

There are several important security-related elements to consider in making cellular phone calls:

- 1) *Service is not assured.* In an emergency and during peak usage periods, call overload can lead to denial-of-service for individual phone calls. Spurious or intentional signals sent by third parties can cause calls to hang up. A moving user can experience dead spots within the service area. In certain locations, such as urban areas, the call coverage can be very spotty due to electronic and physical interference. Transition of calls between cells is not assured. Since cellular systems are implemented based on user population, many areas with low population density may not have cellular service at all.
- 2) *The user is identified.* As soon as a cellular phone is turned on within a service area (a call need not be made), the user is identified to the entire system. The user identification is broadcast within the cell in response to interrogation from the cellular system over-the-air signaling channel.
- 3) *The user location becomes known.* As soon a cellular phone is turned on within a service area (a call need not be made), the location of the user is identified to the entire system. The user is located to within a fraction of the cell area (typically several square miles).
- 4) *The user's information is exposed over the air.* Both the signals transmitted from the user, and the signals from the other party to the call are available over-the-air within the cell site. The equipment required for third parties to intercept calls is inexpensive. Nothing more than a standard cell phone is required to accomplish the interception. There are multiple-hacker web sites that provide information on how to convert a cell phone into an interception-scanning device. The use of high-gain antennas (also cheap and readily available) can extend the interception capability well beyond the cell site itself.
- 5) *An adversary can readily deny service.* Cellular signals can be readily jammed and are subject to interference also. Several vendors make intentional jammers to prevent cell phone operation on a given premises.

- 6) *CDMA technology provides lower signal exposure.* CDMA transmissions are less readily intercepted than TDMA transmissions, but CDMA transmissions are not, by any means, invulnerable.
- 7) *Intelligibility of calls may be poor.* Basic cell phones that use analog user channels can suffer from noise. Digital channels use low data rate voice encoding that can suffer quality loss through conversions from digital to analog and back in the telephone and cellular networks.
- 8) *Users can be spoofed.* Through theft of equipment or reprogramming of IDs, third parties can adopt the identity of a user, and make misrepresented calls.
- 9) *User cellular telephone instruments are vulnerable.* As equipment becomes more sophisticated, more information is stored within the cell phones themselves. Several cellular phone models include a palm top computer as part of the instrument. A stolen cellular instrument may contain much more sensitive information than just the user's identification.

## 5.2.1.7 Framework Guidance

### User Advisory

- Cellular phones are adequate for general-purpose traffic, but are typically unsuited for high reliability requirements. Numerous Government organizations and law enforcement agencies use cellular telephones for general-purpose traffic, but use specialized security devices and private networks, e.g., LMR for critical communications.
- Several cellular providers offer over-the-air encryption of user information, but the security is applied only for the air link, not through the telephone network. In all cases, except for the use of NSA endorsed Type 1 instruments, commercial cellular encryption is not suited for classified information exchange. Discretion in sensitivity of information transmitted is necessary.
- Digital telephone services are somewhat more private than analog systems. Interception of analog conversations is trivial, whereas a small amount of sophistication must be applied to intercept digital connections. Also, digital connections are more readily secured through encryption, should the option be available. Use of digital cellular phones is recommended.
- Use of CDMA technology is preferable to use of TDMA from a signal interception viewpoint.
- Users must protect their cellular phone instruments from theft or loss. The cost of the instrument may be trivial compared to the value of information contained on the instrument.

### Desired Security Solution

- Users within the NII and DII require reliable service with assurance of data integrity and confidentiality, as well as protection from handset cloning and misidentification.

- Any cellular/PCS network should provide over the air security (at a minimum) for both voice and control channel information.
- End-to-end security for user conversations and data transfers is required for U.S. Government sensitive and classified operations.
- Users should be protected from RF attacks and traffic flow analysis attacks.
- Systems should provide capabilities for users to be restricted to absolute need in the use of options available within the systems, e.g., Caller ID, thus minimizing the amount of traffic-related information sent over the air.

## Best Commercially Available Solution

- The best current solutions involve using a PCS phone or a GSM phone with a SIM card to provide user I&A.
- Cellular providers have adopted RF signature evaluation techniques to find stolen cellular user instruments.
- Network providers currently secure billing information through the cellular and PSTN networks.
- GSM standards provide for encryption of user channels within the provider secure infrastructure, i.e., as far as the wired telco interface.

## Technology Gaps

- Adequate security mechanisms to implement Type 1 security for U.S. Government classified operations, for example insertable or software-based high-grade encryption.
- Protocol-sensitive encryption techniques to protect multiple data protocol type.
- SMI within the service provider network to include user security privilege establishment, maintenance, and distribution.
- User-operated control and provisioning systems to allow rapid reconfiguration of user privileges to modify services in emergency quick response operations.
- Modified modulation techniques for spread spectrum systems; e.g., CDMA, can decrease the effect of electronic jamming, and reduce probability of detection for covert users.

## 5.2.2 Low Earth Orbiting (LEO)/Medium Earth Orbiting (MEO) Satellite Telephone Networks

LEO and MEO satellite telephone networks, often referred to as MSS, are the next stage in worldwide, portable telephone connectivity. Unlike the cellular/PCS systems discussed earlier in this section, these handsets will provide telephone connectivity from anywhere in the world that the subscriber elects to pay for service. The traditional cell structure and roaming envi-

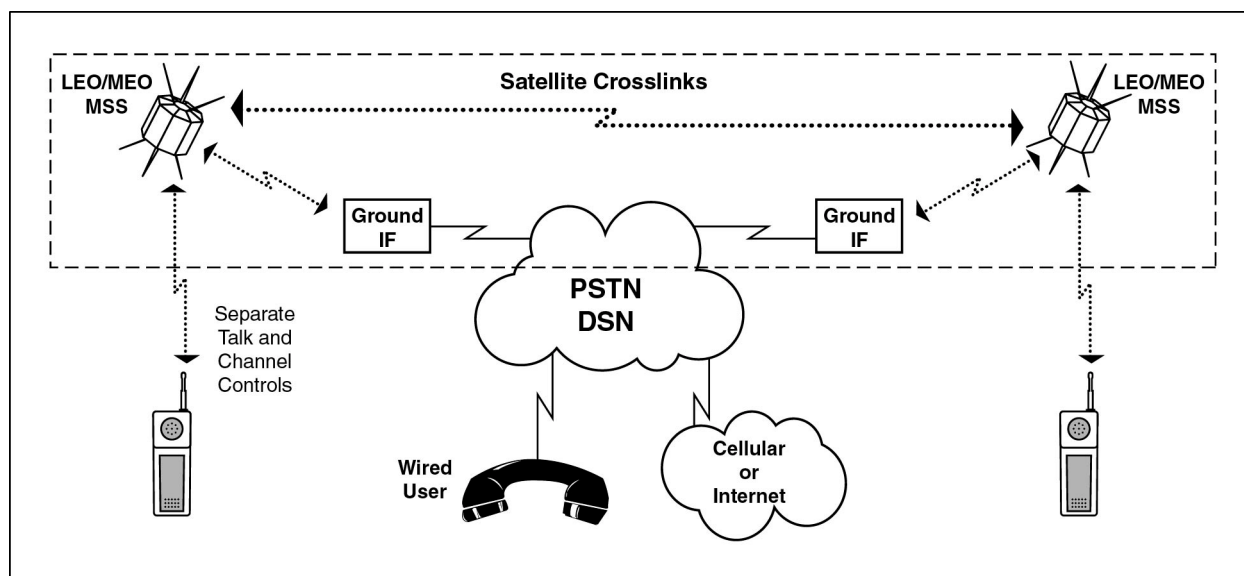
ronment changes significantly because the cells are now moving, and the users are remaining relatively stationary compared to the faster moving LEO satellites.

LEOs circle the planet many times each day at orbit altitudes of 300 to 1,000 miles. The engineering is very complex because these systems cover large areas with many small, low-powered satellites. Currently, only two satellite services are scheduled to be partially available now or in the near future: Iridium and Globalstar. In one case, there will actually be hand-offs between the satellites as shown in Figure 5.2-3. Advantages of these services will include worldwide coverage, the ability to use portable phones and automatic searching for a terrestrial (cellular) service before switching to the satellite. Many MSS phones scheduled for commercial use operate with local digital cellular networks as well as the satellite network. Because of the present high per minute cost of satellite communications, the phone will/should first try to access a local cellular system when making a call. If no cellular service is available, the satellite service is used.

### 5.2.2.1 Target Environment

The target environment is very similar to the cellular case where a user is making or receiving a phone call from a portable mobile user instrument to another portable instrument, to a wired telecommunications user, or to a cellular telephone. In this environment, the user and recipient can be anywhere in the world.

As previously presented for the cellular case, the elements of Figure 5.2-3 can be broken into three major sections: the user environment, the service provider network, and the public network. The user environment consists of the handheld phone and associated user, as well as the talk and control channels. The service provider network infrastructure includes all equipment and connections from the satellites and earth stations, the satellite control infrastructure, and the ground entry points that interface to the PSTN. The public network includes connections to wired users, the Internet, and to other mobile network providers.



**Figure 5.2-3 Mobile Satellite Subscriber Environment**



## 5.2.2.2 Consolidated Requirements

The following requirements are proposed for Government utilization of MSS capabilities.

### 5.2.2.2.1 Functional Requirements

- Global coverage area for call transmission and reception.
- Continuation of call connection from satellite to satellite.
- User and recipient I&A.
- Voice and data confidentiality and data integrity.
- Transmission of voice and data.
- User geolocation capability (both beneficial and a vulnerability).
- Long user instrument lifetime (battery power).
- Accurate and timely billing procedures.

### 5.2.2.2.2 Networking Environments

- Cross-connected satellite constellation for primary call handling (vendor or service provider proprietary protocols).
- Data transmission capabilities up to 19.6 Kbps currently for e-mail and other short message services.
- Interconnection to PSTN, cellular networks, and data networks.
- Worldwide paging services also available through LEO satellite networks.

### 5.2.2.2.3 Interoperability Requirements

- User instruments that can be used with the MSS system and with cellular telephone systems.
- Interfaces with all PSTN systems worldwide.
- Sufficient digital voice quality to traverse the PSTN and be intelligible in cellular systems.

### 5.2.2.2.4 Anticipated Future Requirements

- Increased bandwidth to support data transfer.
- Increased voice quality for conferencing.
- Reduced cost of user instrument to expand availability.
- Support for SMI functions.

## 5.2.2.3 Potential Attacks

### 5.2.2.3.1 Passive

- Largely the same as for cellular RF emissions vulnerabilities.
- Interception of data from the satellite downlink transmission can be accomplished from anywhere in the satellite footprint (larger space than for cellular). The only drawback for the adversary in this case is the volume of information to be processed.

### 5.2.2.3.2 Active

- Denial-of-service attacks by electronic jamming.
- Like cellular systems, network attacks through LEO/MEO satellite systems are somewhat limited in scope. An adversary cannot access the entire telephone network simply by intercepting one telephone call. In other words, local access does not allow universal system access as it would in the case of a LAN connected to the Internet.

### 5.2.2.3.3 Insider

- Modification of handsets before delivery to customer.
- Duplicate handset and user identification information can be loaded into a second phone (non-simultaneous use).
- User location information available to service provider.

## 5.2.2.4 Potential Countermeasures

Many of the countermeasures discussed in the Cellular/PCS section also apply to satellite telephones. Theft of service will most likely be the primary goal of any hacker on the MSS telephone network. Theft of information and eavesdropping will likely be a secondary concern for providers, but will be critical to certain Government users. Service providers must ensure control channel information is secure, and procedures must be in place to provide user I&A in order to prevent theft of service. Providers must also permit the use of end-to-end confidentiality mechanisms to protect user information.

With a cellular structure, creating some type of SMI incorporating key management and other countermeasures is easier within a country. Any SMI used in the LEO network must fit into more of a global management structure. However, as costs drop and satellite telephony becomes more popular, usage by customers within both the DII and NII will likely increase. Before these telephones become useful for customers in the DII transmitting sensitive information, sufficient countermeasures must be implemented to provide privacy, authentication, and message integrity in accordance with the level of information being transmitted.

Use of some sort of token or smart card with the telephone handsets can also be integrated into the satellite network. As with cellular systems, SIM cards may provide the best countermea-

sure to enable user authentication and key management. Only authorized users would be able to access the satellite network. Also, if a phone is stolen, the user can notify the service provider, who then de-activates the SIM card in the stolen phone. The phone can even be programmed to flash “Stolen Handset” to notify the thief that the handset is useless.

## 5.2.2.5 Technology Assessment

As of this writing, service has been initiated on both the Iridium and Globalstar networks. Proposed technologies include dual mode (GSM/MSS) handsets, voice and data transmission, paging, facsimile and position location. Iridium will use a combination of Frequency Division Multiple Access (FDMA) and TDMA multiple access technologies, while Globalstar uses CDMA. Development of a Type 1 secure handset for end-to-end confidentiality in the Iridium network is under study.

### Usage Cases

As stated for cellular usage cases, other sections of this Framework have addressed several cases involving connecting equipment at one classification level to equipment at the same, or different classification level across both trusted and untrusted networks. These cases are clearly an IATF issue, and also apply in the MSS domain. In the case of wireless communications, all transmissions can be thought of as connecting to an unknown environment because of the nature of RF transmissions and the ease of signal intercept. Thus, the descriptions of each of the specific cases addressed in this Framework remain unchanged for the wireless environment.

The sample case of an MSS call can be treated in a very similar manner to that of a cellular call scenario described earlier. If we take the earlier cellular case of calls to another MSS telephone, a wireline-connected standard telephone, or a cellular telephone, the cellular vulnerabilities presented in Section 5.2.1.6 (Usage Cases) exist with some modifications as described below:

- 1) In most cases, the MSS user must pre-register with the service provider for specific “roaming” access areas outside of home territory.
- 2) The extended satellite footprint makes user information more available to interception since the terrestrial range over which the RF signal is broadcast is on the order of several hundred miles.
- 3) For at least one MSS service (i.e., Iridium), user coverage is global. In other cases (e.g., Globalstar/ICO), far north and south latitudes are not covered.
- 4) Transmission rates are typically lower for MSS services than for cellular services. Since digital voice rates are reduced, voice quality is reduced. Connections across MSS and cellular systems may suffer degradation in voice quality to the point where user voice recognition is not possible.

## 5.2.2.6 Framework Guidance

### User Advisory

- The risks for users in using MSS services are similar to those for cellular. The range of interception for MSS calls is increased, but the risk of geolocation is reduced. Keep messages short for both security and financial reasons.
- There is insufficient data concerning the operability of MSS systems to make definitive statements on system availability and loading. Request provider information on call completion rates.
- The development of instruments and protocols for high-grade end-to-end confidentiality are under study. If you are addressing user requirements for your organization, contact NSA for status of efforts.

### Desired Security Solution

Ideally, a MSS telecommunications network will provide confidentiality for both talk channel and control channel information. Users within the Government require reliable service with some assurance of data integrity and confidentiality, as well as protection from spoofing and misidentification, e.g., handset cloning. An integration of the smart card technology used in GSM phones with the satellite phone handsets could help provide adequate protection for users.

### Best Commercially Available Solution

Currently, only the Globalstar networks is scheduled to be operational by the end of this century. Other than methods to secure control channel information, no security solutions have been integrated into this system. The primary security needs for satellite telephone services are end-to-end confidentiality for user information, and the protection of caller and calling party identification.

### Technology Gaps

- Adequate security mechanisms to implement Type 1 or Type 2 security.
- SMI within the service provider network.
- Protection of stored information in user instruments.
- As wireless telephones increase in complexity and become more like personal computers, user handsets will require a way to provide secure data storage using SIM cards or other types of token.

## 5.2.3 Wireless Local Area Network (WLAN)

WLANs are quickly gaining popularity in multi-user environments. A WLAN can be used as a stand-alone network, or as is most often the case, it can be used to increase the range, flexibility, and user mobility of a larger network. WLANs are typically implemented with PC cards inserted into network processors, and can also be implemented in portable devices such as hand held computers. A WLAN uses the same transmission (Ethernet is typical) and data protocols, e.g., IP, as its wired equivalent but provides a lower bandwidth, e.g., 2 Mbps versus 10 Mbps or 100 Mbps for Ethernet. The typical implementation for RF communications is a collision avoidance direct sequence spread spectrum or frequency-hopped protocol under the Institute of Electrical and Electronics Engineers (IEEE) 802.11 standard. Members of a WLAN communicate one at a time as on an Ethernet rather than in an overlay of signals as occurs in CDMA cellular systems. Multiple WLAN nets can then be overlaid in the same location and frequency range by using different spreading or hopping sequences. WLAN members have a connection distance measured in the range of 100 to 1000 of feet depending on the environment, e.g., office building, and open space.

WLANs have gained entrance into the marketplace primarily in the vertical markets of health-care, retail, manufacturing, warehousing, and academia. These markets have leveraged from the productivity gains of using hand-held terminals and notebook computers to transmit real-time information to centralized hosts for processing. Primarily WLANs are an advantage when mobility, scalability, and installation speed, simplicity, and flexibility are important requirements. An interesting example of a large-scale WLAN integration is the Fox Tower building in Portland OR. The Fox Tower will feature connectivity to a high speed fiber-optic network, including satellite transmission, digital phones lines, WLANs, video and high speed digital subscriber line access to every tenant on every floor, regardless of each tenant's current technology capacity. This is an example of the architecture providing information technology infrastructure in a flexible, scalable plan to minimize the cost of constantly upgrading the system infrastructure as tenants move or change technology.

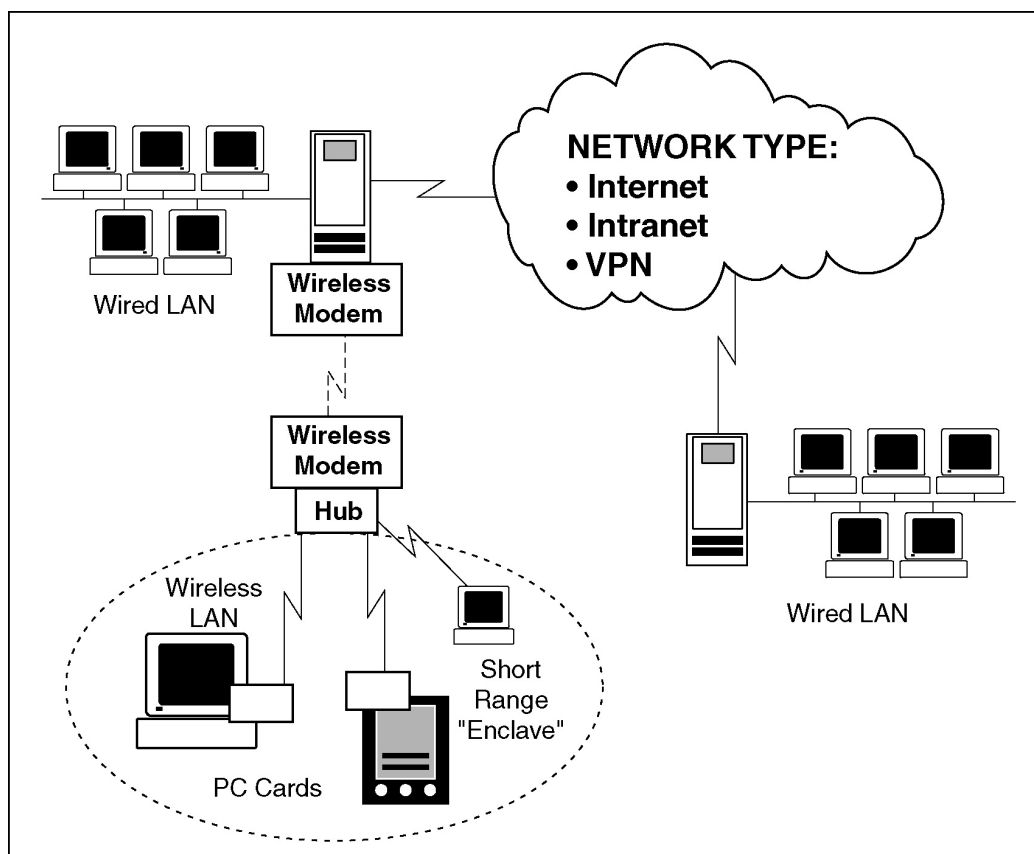
### 5.2.3.1 Target Environment

The WLAN provides flexibility for movement of net members but requires a high degree of collocation of the wireless segments (communications range on the order of 500 feet). WLANs are often used in offices and facilities where the wiring required for a standard network has not been installed. Other applications include provision of network interconnection where the nets must be configured and torn down rapidly. A tactical military command post or forward air base is an example of the latter. The target environment, shown in Figure 5.2-4, has been drawn to represent the case where a WLAN extends an existing network through a wireless modem link.

The WLAN environment is a notable exception to the definition of “wireless” provided in Section 5.2 (Wireless Network Security Framework) in that, in the WLAN case, the user owns the wireless infrastructure (however small that may be). The user buys the components and does not need to rely on a service provider for WLAN operation. This fact provides flexibility in location, mobility, and applications. However, the WLAN is tied to a wired LAN environment in most cases, thus reintroducing “borrowed” infrastructure requirements.

The wired infrastructure to which the WLAN is connected can be formulated in several ways. As shown in Figure 5.2-4, the “cloud” can be the Internet, or a secured environment composed of an Intranet or VPN. The security implications of connecting WLAN components to an Intranet or VPN are of particular importance. It must be noted that the range from which an “observer” can observe (detect or read) the signals emanated from the wireless connection is always greater than the range over which the WLAN will operate. Very simply, the use of high gain directional antennas from a remote location provides the same receive signal strength that can be achieved by a close-in user with a standard antenna and receiver.

The key elements of the environment are the physical space where the WLAN is implemented (size and type of physical environment and its perimeter), the level of classification or sensitivity of information handled in the system, and, as mentioned in the previous paragraph, the wired interconnect mechanism. Special cases of high to low classification, firewalls and other wired LAN security elements are assumed to be handled by the wired LAN segment of the target environment.



**Figure 5.2-4 WLAN Environment**

## 5.2.3.2 Consolidated Requirements

Users of WLANs typically connect through an access point to a larger wired network. Each access point can represent a separate user domain, or multiple access points can be assigned to the same domain to increase data throughput in high usage areas. When connecting a WLAN to an existing network, system administrators must be careful not to weaken the existing network security of the wired LAN. The use of VPNs, as discussed in Section 5.3 (System High Interconnections and Virtual Private Networks (VPNs)), will play a large part in ensuring adequate security for WLANs. Without access controls at the wireless nodes, an attacker can gain universal access to the entire network by simply penetrating a single node.

Additionally, a distinction must be made between use of a WLAN in a standard office environment versus a highly mobile or tactical environment. An office environment will typically require a network to handle higher traffic loads and a large number of users. Tactical environments, on the other hand, will usually operate in a hostile environment. Traffic loads may vary, and networks will typically consist of fewer, more mobile users than wired cases. Requirements may differ dramatically between the two environments. The following is a list of proposed requirements.

### 5.2.3.2.1 Functional Requirements

#### User/Mobile Terminals

- Provide access control for restricted domains.
- Provide user I&A mechanism.
- Ensure VPN software compatibility to support data confidentiality.

#### Access Points/Network Equipment and Configuration

- Strong access control.
- Ensure network bandwidth availability. The network must be fast enough and able to handle a large number of nodes without becoming unusable.
- Ensure data integrity.
- Provide continuous authentication of all users connected to a WLAN.
- Establish secure wireless domains for each access point.

### 5.2.3.2.2 Networking Environments

- Ability to communicate with wired networks through a wireless access point within range of the LAN at data rates sufficiently high to prevent congestion.
- Ability to communicate at close range among mobile elements (ad hoc network) as in a field tactical situation.
- Provision of spreading codes that minimize recognizable transmission characteristics for Low Probability of Intercept/Low Probability of Detection (LPI/LPD) capability.

### 5.2.3.2.3 Interoperability Requirements

- Interoperability between networks operating at different data rates, e.g., WLAN at 2 Mbps and wired LAN at 100 Mbps; without adversely affecting quality of service.
- Networks using different modulation schemes cannot communicate directly with each other without any conversion. Both Direct Sequence Spread Spectrum (DSSS) and Frequency Hopped Spread Spectrum (FHSS) modulation are part of the IEEE 802.11 WLAN standard. In the standard network environment, gateways are used to translate between networks from one protocol to another.
- Co-locating WLAN systems must not cause interference problems with other wireless systems in the vicinity. Spread spectrum modulation attempts to minimize this interference. However, with the common 11-bit spreading code, WLAN systems will not attain a processing gain much higher than 10 dB (Federal Communications Commission (FCC) minimum). Longer spreading codes would increase processing gain and could improve data security.

### 5.2.3.2.4 Anticipated Future Requirements

- Wireless networks must allow for the evolution and re-configuration of the network and associated components without disruption of service.
- Higher data rates will likely lead to more frequent transmission of time sensitive data such as audio and video files. Current standard data rates of 1 or 2 Mbps are far too slow for practical video transmission given that a multi-user LAN begins to saturate at an aggregate throughput of approximately 10 percent of rated speed. Also, transmission of large text or image files can cause congestion in a WLAN. WLAN data rates are quickly approaching 10 Mbps. However, even these higher data rates only allow for 6 to 7 Mbps throughput. When encryption is added, throughput drops even further. An active office environment with multiple users would provide considerable strain on a LAN at these speeds. Future requirements will likely call for throughput rates near or above 10 Mbps.
- Current WLANs can optionally apply low grade data scrambling or basic encryption to the transmitted data. All the header information is frequently sent over the air in the clear. This causes weak traffic flow security, a problem that will be discussed in the Potential Attacks section below.
- If WLANs are to be used in a classified environment, individual node identity and message header information may be classified and thus will need to be protected at a higher level of security than presently available. This will require capabilities akin to the Network Encryption System (NES) or other robust encryption discussed in Section 5.3.5 of the IATF, but with a portable form factor.

### 5.2.3.3 Potential Attacks

A WLAN without appropriate security mechanisms in place can add critical vulnerabilities to a network, making it easy for an attacker to penetrate. With WLANs, an adversary no longer re-



quires physical access to the network, as in a wired situation, in order to exploit a wireless system. This physical access is particularly important to an adversary in the case of VPNs and Intranets where physical access is required if those systems are properly established and protected in accordance with the IATF recommendations. Addition of a WLAN to a VPN or Intranet removes the physical access requirement for an adversary to penetrate the system.

### 5.2.3.3.1 Passive

- Signal detection and intercept are readily accomplished due to the limited requirements for diversity in spread spectrum systems. The standards are public in IEEE 802.11, facilitating signal detection.
- WLAN signals are designed to penetrate office walls and maintain user connectivity at significant distances up to several hundred feet. Therefore, an attacker has the advantage of operating without requiring access to a protected facility, and the attacker can use high gain antennas and receiver equipment to recover a signal. (Note that this is a major difference from a wired architecture. While some devices on a wired network may inadvertently radiate, they are not designed to do so. Cable shielding and the use of fiber optic cable for network connections make it difficult for an adversary to tap on to a wired network without gaining access to the actual cabling.).
- A passive attacker can determine critical information about network architecture just by monitoring message headers, even if all the transmitted data has been encrypted. While this may be acceptable for Government and some DoD applications, many Government sensitive networks and military tactical networks would prefer not to divulge critical information about network nodes. Therefore, there is a clear requirement for inclusion of strong message confidentiality, and good traffic flow security (packet header cover) in future WLAN designs.
- In the case of a wireless network attached to a VPN or Intranet, the RF link opens access to an eavesdropper thus bypassing any Firewall or encryption security used on the wired network. In the case of a network connected to the Internet, access is more readily achieved using the wired network, so the RF passive attack is unlikely.

### 5.2.3.3.2 Active

- Attacks on a WLAN can be accomplished easily with the proper network analysis equipment. Standard network sniffers can be adapted to analyze wireless network packets. Current sniffer technology allows the sniffer software to be run from a laptop computer.
- Denial-of-service attacks, though not specifically network based, can have drastic effects on critical DII and NII networks if not properly detected. WLANs operate like any other radio in that the receiver must maintain an adequate signal-to-noise ratio in order to maintain a link. When the noise overpowers the signal and any processing gain, proper reception will not happen. If an adversary decides to jam an access point or a major portion of the wireless network, the WLAN will not continue to function. However, this type of attack, and the source of the interference, would be easy to detect and correct. On the other hand, if an attacker directs a jamming signal at only one node, the rest of

the network has no way of knowing why that node has gone down. In fact, many of the access points, i.e., wireless hubs, on the market today will continue to show a valid connection to that node even if it is currently unreachable. If a WLAN is used in a critical part of the NII, preventing denial-of-service attacks will be a major issue to address.

- Network information available to an adversary can lead to spoofing attacks using directional transmission aimed at the system RF hub or at single node. The attack against a single node is more difficult to defend against because the RF hub would be unaware of the interference.

### 5.2.3.3.3 Insider

- An insider on a WLAN can often have access to access point configuration files. Without proper administrator authentication procedures at the access point, a user can modify these configuration files to increase the vulnerability of the entire network. For example, access points will usually only forward a message to their wireless nodes if the intended recipient is in that accessed point's domain. Thus, the wireless link is more efficient, and an attacker cannot easily view messages between nodes on the wired network. A malicious insider could modify the access point configuration to pass all or none of the network messages on to its nodes, if proper administrative authentication procedures are not in place.
- Similar to a wired network, many insider attacks are available in a WLAN. While user privileges can be set on a network server by the system administrator, there is no mechanism in place to prevent a legitimate user on the system from entering more private areas on the network. File privileges can be set on sensitive files, but if a privileged user wants to take advantage of a WLAN, there is no mechanism to prevent this. Again, this problem is not specific to wireless networks and was addressed in earlier sections of the Framework.

### 5.2.3.3.4 Distribution

Hardware or software modification in transit could be used as a first step in a complete attack by which an adversary eventually causes the system to send data or allow access by way of electronic connections to information for which he is not authorized. These attacks are more readily prevented using physical and operational security techniques, and are not a primary emphasis in this section.

## 5.2.3.4 Potential Countermeasures

Many of the countermeasures used in a wired network, and those described in Section 5.3.4 (Potential Countermeasures) for VPNs, also apply to the wireless case. In general, maintaining privacy is accomplished by appropriate use of confidentiality mechanisms. If a WLAN is employed in a classified application, the strength of confidentiality mechanisms must be sufficient for withstanding national laboratory strength attacks.

As discussed in the Section 5.2.3.3.1 (Passive), traffic flow security is a major issue. Unfortunately, a WLAN cannot simply implement a constant bit rate leased line or other traffic shaping

mechanisms. Leased lines in the wireless case do not apply, and traffic shaping may severely limit the throughput of the wireless link and interfere with the collision avoidance mechanisms in place. One way to provide some traffic flow security would be to route all wireless traffic through secure tunnels.

Wireless network sniffers used in conjunction with bit generators can be used to insert messages into a wireless network that appear to have originated in the network. Continuously authenticated channels can prevent insertion of information into the channel that can lead to short plaintext attacks that allow cryptanalysis by guessing known responses to known short messages.

Prevention of Denial-Of-Service (DOS) attacks is a difficult issue. Although, in some respects the wireless case is very much the same as a denial-of-service attack on a wired network. Network administrators must implement proper authentication software to prevent the manipulation of network hardware. In the wireless case, simple signal detection mechanisms can probably detect and locate an obvious RF jamming signal as easily as an administrator on a wired network could detect a broad DOS attack.

### 5.2.3.5 Technology Assessment

The technologies for WLANs are targeted at minimized bandwidth licensing requirements. Since users own their system infrastructure for WLANs, the low power and spread spectrum techniques that support non-licensing of the spectrum are valuable to the user community. However, users, particularly Government and DoD users, are cautioned that unlicensed bandwidth in the U.S., e.g., 2.4 GHz band, may require licensing for use in foreign countries. Federal licensing authorities must be consulted on foreign requirements for bandwidth and spectrum allocation before systems are implemented in foreign countries.

Frequency hopping and direct sequence spread spectrum are both defined in IEEE 802.11 for WLAN applications, and both have been implemented by product vendors, but DSSS is the more popular implementation. Limited LPD is provided by the waveforms, but the 802.11 standard is sufficiently restricted in spreading patterns, that such protection cannot be deemed suitable in military environments. The Anti-Jam (AJ) protection that is afforded is similarly weak for the same reason.

Current encryption and data scrambling methods used in WLANs provide minimal data protection, and are not suitable for protection of classified information. The data encryption techniques for commercial WLAN are insufficient for other than privacy. Presently key lengths are restricted to 40 bits. The casual probe will not achieve access, but the strength of the cryptography will not withstand a more determined attack. Cryptography that provides security for transfer of header information is not in place, and is not easy to implement. DoD products such as TACLANE cryptography are available for high-grade protection of over-the-air signals. Development of PC card-based Type 1 security devices is also under study. The interfaces are complicated by use of such products because the commercial capabilities are meant to plug directly into processing elements. The DoD cryptography must be inserted between the processing and transmission elements. The TACLANE is transportable, but not man portable.

Operating frequencies vary according to product vendor and system. Presently, the 2.4 GHz band is the most popular, however, higher data rates are achieved with larger bandwidth in the

5.6 GHz range. It has been found in certain applications environments that interference problems can occur. Notably, microwave ovens have been found to “jam” some WLAN systems. The RF technologies used in the GHz range communications systems include antennas that vary from 2-3 db isotropic to directional gains in excess of 20 db. In fixed plant configurations (or portable configurations that remain in one location during operation), the directional antennas can be used for nodes of a WLAN to increase range to a distance of several miles. Such nodes cannot then be highly mobile, since directional antennas must aimed for effective operation. Unfortunately the same antennas can be used by an adversary to expand his probe range to a similar distance.

The wireless modem shown in Figure 5.2-4 provides the capabilities of a microwave transmission system at a small fraction of the cost. The use of such modems, as in the case of microwave links, can readily be equipped with over-the-air confidentiality be applied to the to the modem point-to-point connection. Since the connection is point-to-point, and independent of protocol, there are straightforward solutions provided by commercial vendors and the DoD to provide link encryption security at the requisite security levels.

### 5.2.3.6 Usage Cases

Other sections of this Framework have addressed several cases involving connecting equipment at one classification level to equipment at the same, or different classification level across both trusted and untrusted networks. These cases are clearly an IATF issue, and also apply in the WLAN domain. In general, some level of communications security is recommended for any equipment where there is a connection to a potentially hostile or unknown environment. In the case of wireless communications, all transmissions can be thought of as connecting to an unknown environment because of the nature of RF transmissions and the ease of signal intercept. Thus, the descriptions of each of the specific cases addressed in this Framework remain unchanged for the wireless environment.

As mentioned previously, the type of network to which a WLAN is connected has substantial impact on vulnerabilities, attack approaches, and the damage that can be done. There are three interconnection possibilities in the scenario presented here for WLAN:

- 1) Users connected to a standalone WLAN.
- 2) Users connected to a WLAN that is interfaced to a wired VPN or Intranet.
- 3) Users connected to a WLAN that is connected to the Internet.

Figure 5.2-4 shows the three scenarios. The following security related elements apply:

- 1) *Over-the-air exposure exists.* Although spread spectrum techniques are used, the spreading techniques are public and the signals are not difficult to intercept.
- 2) *Detection range of WLAN signals is much greater than communications range.* Typical WLANs use small omni-directional antennas. High gain directional antennas can pick up signals at much greater ranges than those used for communications (the range can be several miles).
- 3) *Information on any WLAN connected network is exposed.* All communications on a WLAN are exposed to interception. Information on wired LANs to which the WLAN is con-

nected is also exposed to interception. In the case of VPN or Intranet connections, the protective mechanism of those networks may be defeated.

- 4) *IP headers are subject to traffic analysis.* The interception of IP traffic can compromise more than user data through the use of source/destination analysis.
- 5) *WLAN signals can be spoofed.* Just as in the Internet, adversaries can use RF signal paths to masquerade as valid users or to deliver spurious messages.
- 6) *WLANs can be jammed.* Multiple jamming techniques exist for denying service to WLAN users.
- 7) *Low data rates of WLAN segment may reduce availability.* When a WLAN is connected to a high speed wired LAN, WLAN users may experience reduced system availability and grade of service.
- 8) *Service may not be available in mobile systems.* If the WLAN network is developed using mobile components, nulls in signal may exist and users may periodically move out of range of other users or of network access points.

## 5.2.3.7 Framework Guidance

### User Advisory

- As discussed in Section 6.2.6 (Cases) of Remote Access, Top Secret and compartmented information on wireless networks presents extreme risk, and should be handled on a case-by-case basis.
- Do not assume that either the spread spectrum techniques used, or the short communications range of the WLAN components affords any protection against signal and data interception.
- Do not develop standard timing structures for transmissions. Asynchronous operations are preferred. Noise can alternate with real data.
- Use “ping” signals to test channel availability before commencing transmission.
- Do not process classified information on a WLAN without Type 1 encryption.

### Desired Security Solution

- Secure data and header information in sensitive transmissions.
- Provide LPI/LPD of LAN transmissions for tactical situations.
- Protect wireless network against traffic flow analysis through RF transmission patterns.
- Continuously authenticate WLAN nodes to the “parent” system.

### Best Commercially Available Solution

- Personal Computer (PC) Card/FORTEZZA® Card software encryption of data prior to transmission.

- Most manufactures use the 11-bit spreading codes called for in the IEEE 802.11 specifications. However, some manufacturers have modified the selection of spreading codes by implementing a way to select a different spreading code for each transmitted symbol. Thus, an additional level of transmission security is provided.
- The RF protocol, using direct spreading, is provided to increase bandwidth, make use of unlicensed spectrum, and to increase numbers of users that can be accommodated. The same technology also provides a degree of LPD protection.

## Technology Gaps

- Inexpensive, portable packet encryption techniques (mechanisms similar to NES/ Internet Protocol Security (IPSec) functions) for WLANs are not available.
- Increased throughput of at least 10 Mbps is not yet available.
- Improved spreading and/or hopping characteristics of spread spectrum transmissions could be implemented, but are not accommodated in the standards.

### 5.2.4 Paging (One-way and Two-way)

Paging is defined as a broadcast or a duplex (that is, one-way or two-way) communication of short messages to highly mobile users in an area where system infrastructure is available for line-of-sight transmission of the messages. Paging was originally a one-way service provided over licensed channels for delivery of numeric messages. Today, paging can be one-way or two-way and numeric, text, or voice, so users may receive and send multiple types of short messages to/from their portable devices.

Paging can be accomplished over many networks, such as digital cellular, PCS, packet radio, and trunked radio. References to paging in this section apply to the transmission of many types of data over many types of system infrastructure depending on the facilities available to the service provider. Wireless communications providers have entered the paging market to enhance revenue for unused bandwidth in their cellular systems. Paging messages are broadcast when channels are tied up with circuit switched cellular calls.

Pagers have gained widespread market penetration, and they are currently used by a large number of customers in the Government, business, and personal environments. Although paging functions have been integrated into many types of mobile user systems (primarily cellular), paging is expected to exist as a standalone service well into the future because of the low cost of the service and the miniaturization of the user devices. Purely numeric paging will drop in usage, but bi-directional short message service will take up the slack. One industry leader predicts a US paging market of 70 million devices by the year 2005.

From a security and availability perspective, service provider advertising has not painted a totally accurate picture. Because each pager is identified by its own individual “cap code”, and the services are largely digital, there is a perception of message confidentiality. As presented in the news media, DoD and other Federal Government users have frequently become targets of pager attacks in the past. Paging is, in fact, a favorite “easy pickings” target of hackers. Primarily, the attacks have caused only embarrassment to the target organizations, but sensitive

information has been involved in several cases, e.g., the location and plans of Secret Service personnel on a presidential protection mission in 1997.

In paging systems, message delivery is not guaranteed, but is largely reliable. Paging systems are designated as one-way, 1.5, 1.75, and two-way. The intermediate numbers roughly describe the ability of the user device to respond back to the messages and prompts. In general, the paging system does not know the location of a user, so the message is flood routed to all areas for which the user has paid for service, thus increasing message exposure. Pagers above the one-way level are able to identify themselves to the system infrastructure so that the paging message is broadcast more selectively. The selective capability is increasing as more systems provide two way paging. However, the basic low cost service provided by most purely paging vendors is of the one-way variety. Battery lifetime is also a concern from an availability viewpoint; the more complex the device, the shorter the battery life.

### 5.2.4.1 Target Environment

Pagers are used in a wide variety of environments, primarily personal and business, but also for urban police operations, emergency operation broadcasts, and even for White House Secret Service communications. Two-way paging networks can be used by police in their vehicles for preliminary checks of criminal records or to perform quick drivers license checks. Emergency operation broadcasts are used in both civilian and military environments to inform staff of the need to contact authorities. In these situations, guaranteed message delivery becomes critical, while security requirements will vary by users and particular situations. The following requirement list covers many different paging environments and will not apply to every situation.

A generalized environment for pager communications is shown in Figure 5.2-5. This environment is largely available in areas with high population density, since service providers wish to maximize the number of customers for a given (often sizeable) system infrastructure investment. The figure represents cellular towers as the transmission mechanism, but this is not necessarily the case. Paging providers will often rent space for their transmitters on cellular towers (and cellular providers do use the cellular transmission media for paging), but pure paging systems use different transmitters and substantially higher power output due to the restriction of receive sensitivity on miniaturized cellular receivers.

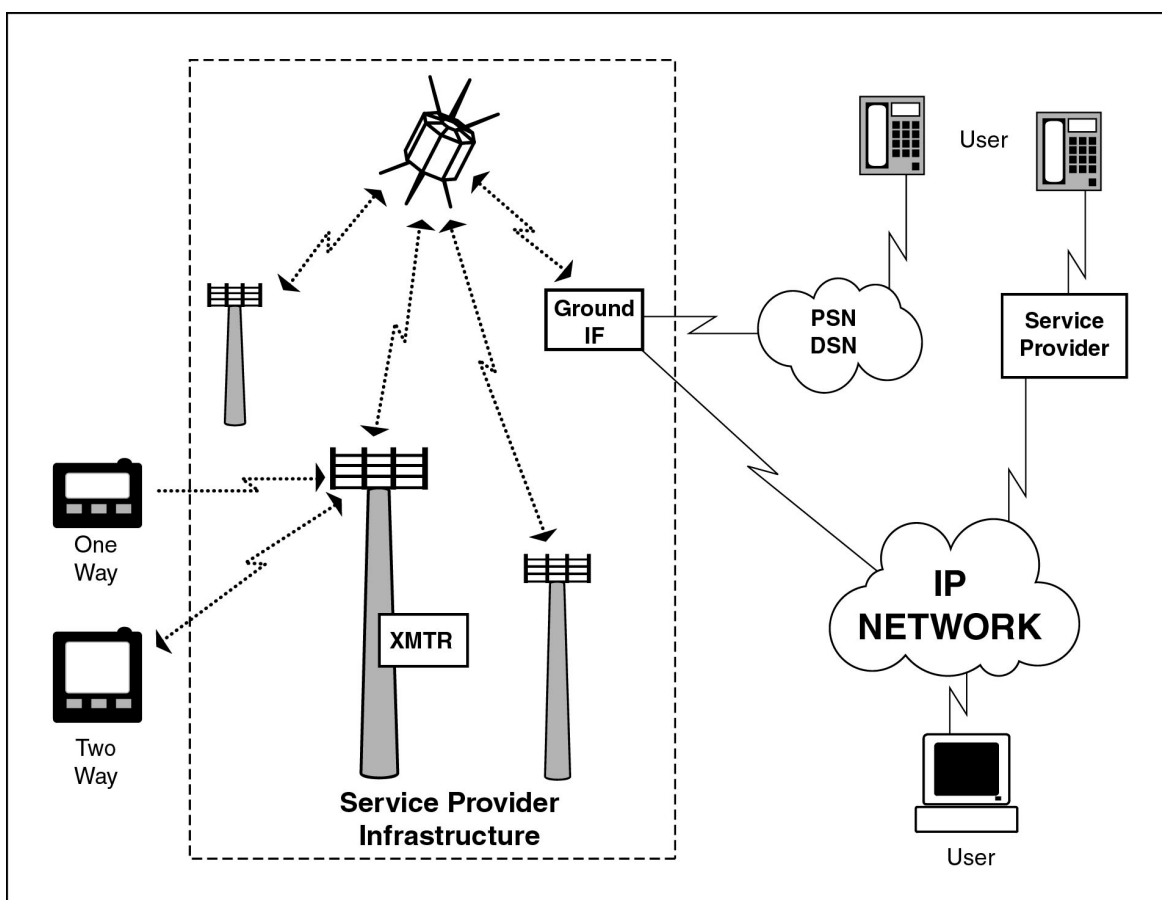


Figure 5.2-5 Pager Environment

## 5.2.4.2 Consolidated Requirements

The proposed requirements for paging operation are varied. The following list represents a consolidated set of functional capabilities that an advanced paging user would find useful.

### 5.2.4.2.1 Functional Requirements

- Receive telephone call-back (numeric) messages.
- Receive short text messages.
- Receive short voice messages similar to voice mail.
- Transmit short messages (numeric, text, and voice) (two-way paging).
- Provide message receipt verification to sender.
- Provide guaranteed delivery.
- Simulcast (reach multiple recipients with a single message).
- Provide confidentiality for message addresses.



- Provide confidentiality for message content over the air.
- Provide confidentiality for addresses and message content within service provider system.
- Provide indication of message receipt on mobile user device.

#### 5.2.4.2.2 Networking Environments

- Both manual and automated interfaces, e.g., dial PIN and callback number, should be available at the service provider for numeric paging.
- Service providers require PSTN interfaces for message initiation.
- Various trunk (bulk transmission) media are required for distribution of messages to the over-the-air transmission sites. These can include leased satellite (as shown in Figure 5.2-5) or various land line or microwave systems (typically leased bulk data services where the provider is only concerned with delivery at the endpoints, and not the distribution path).
- The paging company/service provider requires an interface with the Internet for individuals to send messages to pager customers. Pagers interface with the Internet primarily for send and receive short messages and e-mail. Other web services such as traditional browsing and file transfer are very costly because the user is charged by the number of characters downloaded every month. Pagers must maintain an emphasis on short messages to remain an affordable service.

#### 5.2.4.2.3 Interoperability Requirements

- As paging technologies progress, older paging protocols are slowly decreasing in use. However, there is still a requirement for interoperability with older protocols like POCSAG.
- The Flex protocol has begun to dominate the market in the U.S. Two-way paging protocols like the Motorola Reflex and Inflexion protocols are becoming de facto standards.

#### 5.2.4.2.4 Anticipated Future Requirements

- Provide confidentiality as a for-fee service element.
- Provide authentication of user to enable access to portable paging device.
- Provide authentication of message initiator.
- Increase message storage capacity of user paging devices.
- Provide interfaces to Virtual Private Networks (VPN).
- Provide over-the-air SMI capabilities to include user identification and key management to support confidentiality.
- Provide e-mail filtering and other message related applications.
- Provide interoperability with LEO satellite paging networks for global coverage.

- Provide interfaces to other user devices, e.g., palmtops, PCs, for message transfer and information synchronization.

### 5.2.4.3 Potential Attacks

Pager users often do not consider the possibility that their communications might be intercepted by an eavesdropper. However, eavesdropping on pager traffic is relatively easy to do. Any individual with access to the Internet can download software and instructions on how to intercept pager traffic. Also, lists of pager cap codes, and often PINs, are published for all to see. There is a question of how sensitive the traffic sent over the paging network truly is. Traditional numeric paging simply alerts the paging customer to call a certain number. However, with the advent of text, message, and voice paging, more significant privacy and security concerns exist.

#### 5.2.4.3.1 Passive

- Intercepting pager traffic is readily accomplished, although illegal. Techniques, methods, and suggested equipment lists are posted on the Internet for any individual to read. Message traffic may be broadcast far beyond the area where the intended recipient is located due to the flood routing algorithms used.
- Cap codes and PINs are often sent over the air to new users. An adversary can reprogram a second pager to receive all messages intended for a specific pager without being detected.

#### 5.2.4.3.2 Active

- Attacks on e-mail and messages sent by Internet users are vulnerable to attack as described in earlier sections of this IATF.
- Denial-of-service attacks through electronic jamming of the paging network in a localized area may go undetected by users.
- Spoofing techniques can be used by an adversary to send a message that appears to originate from a different location than it actually does. Without a way to validate message origin, recipients cannot be sure if they have received a valid message.

#### 5.2.4.3.3 Insider

- An insider is anyone having access to a paging service provider's database, customer personal account information, or paging equipment, whether or not this access is authorized by policy. These attacks could be motivated by deliberate malice, or could be the result of unintentional mistakes on behalf of the user or service provider. Results of a deliberate attack can be especially damaging to the organization's information system due to the attacker's access to the information, their advantage of knowing the network's configuration, and thus their capability to exploit the network's vulnerabilities.

- A second type of insider attack involves theft of service or equipment by service provider representatives.

#### 5.2.4.4 Potential Countermeasures

- Users must be educated as to the capabilities and vulnerabilities of their pager service.
- Encryption methods can be provided for message confidentiality (net or public key).
- Authentication methods for both message initiators and recipients can be provided.
- Guarantee of delivery can be provided through use of 1.5, 1.75, and two-way paging techniques.
- Anti-Jam (AJ) and Low Probability of Intercept (LPI) communications techniques.

#### 5.2.4.5 Technology Assessment

Since pagers are dependent on the RF media for message delivery, over-the-air confidentiality is a primary concern. Present packet structures for paging messages provide very little message bandwidth (on the order of dozens of bytes for older systems and hundreds of bytes for advanced paging systems). Additionally, most providers charge for their service by the byte delivered. The narrow available bandwidth creates difficulty with the overhead that is introduced for secure message delivery. Such overhead includes key distribution, synchronization, and reformatting of messages, e.g., Uencoding, for delivery over packetized networks. New technologies are continually increasing the bandwidth available to pager systems, so overhead concerns will be reduced.

One vendor has developed a pager security technique that employs over-the-air encryption and firewall wired network access. Although promising, the technique does not provide confidentiality in parts of the service provider system infrastructure.

Pagers presently have minimal storage and programming capacity to support security mechanisms. Hand held computers and cellular phones that can be programmed or provided with ancillary devices, e.g., PC cards, to provide paging service are candidates for insertion of security mechanisms, but these devices do not fit into the miniature device pager-only scenario.

Guaranteed message delivery remains an issue when a return path is not available. However, procedural methods like telephone callback can be implemented to give assurance of message receipt. In fact, telephones can be busy, and e-mails may not be delivered, so the pager scenario is not necessarily of lower assurance than other message delivery mechanisms. If message assurance is required, then two-way paging techniques can be employed at higher costs than those for one-way service.

The interfaces provided to pager devices are minimal at this point primarily due to cost and size considerations. Off-line security measures (authentication, encryption) can be considered if interfaces are provided for elements such as smart cards or CompactFlash cards. New standards for RF interfaces to miniature devices, e.g., Bluetooth, could more readily support security services.

## 5.2.4.6 Usage Cases

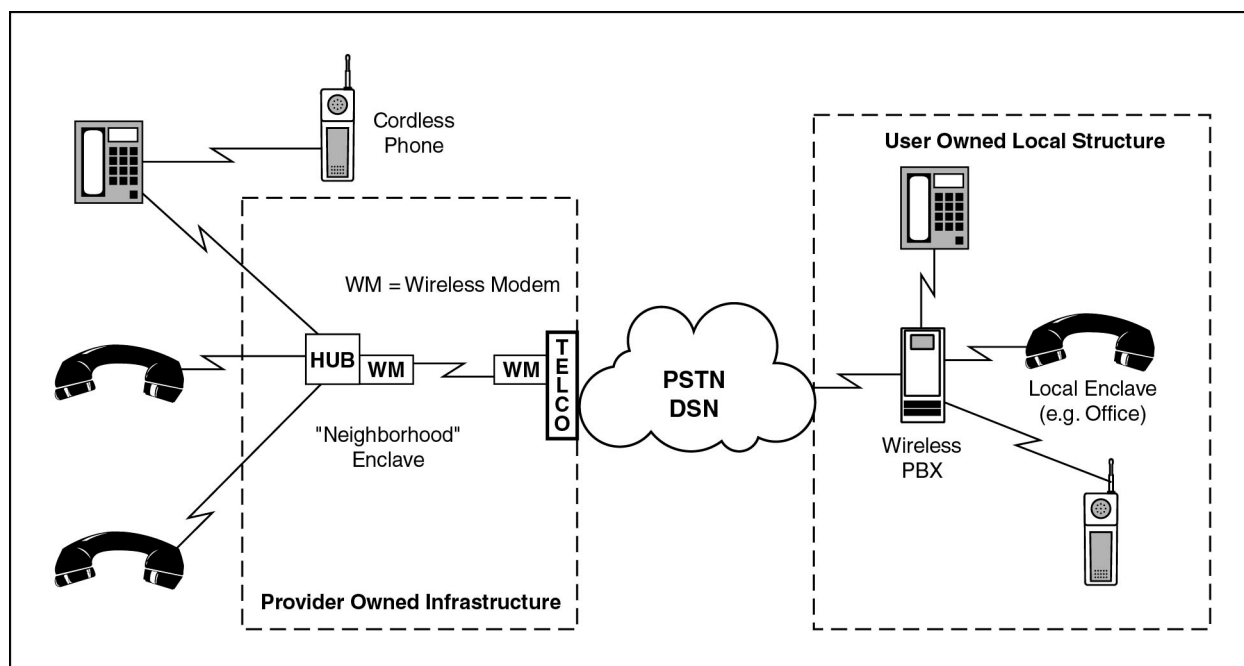
The usage cases for paging involve several different configurations as shown in Figure 5.2-6. The potential use of the Internet, VPNs, or other IP-based network types in the scenario results in vulnerabilities discussed in other sections of this document in dealing with the wired network systems and system infrastructure. However, unlike the WLAN situation, the use of pagers with network connections does not necessarily increase vulnerabilities of the wired network. Pages are sent using a set of pager unique protocols rather than IP protocols. Thus the exposure of the IP network is not as great as would be with a WLAN connection.

As shown in the figure, there are three different access methods for initiation of the pager message.

- 1) Sending party uses Internet to reach service provider.
- 2) Sending party uses standard telephone call to reach service provider.
- 3) Sending party uses cellular telephone to reach service provider.

The page message can be delivered under several scenarios that are service and service provider specific.

- 1) One-way page with no response from the recipient.
- 2) 1.5 or 1.75 way page with limited response back to the provider system from the message recipient.
- 3) Two-way page where specific full message can be developed in response to the pager message.



**Figure 5.2-6 Pager Usage Scenario**

When employing a pager system for sensitive and important messages, the mobile user must be aware of the characteristics of pager transmission.

- 1) *Over-the-air interception of pager signals has a broad range.* Since pager signals are broadcast to the entire coverage area of a pager system, an adversary can intercept messages from anywhere in the pager coverage area. The requirements for interception are trivial and available on many hacker web pages. Also, in one way paging systems, messages are broadcast multiple times to increase probability of delivery.
- 2) *All pager messages pass through an insecure provider network.* The provider may be telco connected, or connected through the Internet.
- 3) *Message delivery is often not guaranteed.* One way pagers do not assure delivery, or at least do not inform the message sender that the page was not delivered.
- 4) *Messages can be stored in low security environments.* Some providers will store messages for later repeated transmission if acknowledgements are not received.

## 5.2.4.7 Framework Guidance

### User Advisory

- Pagers have all of the vulnerabilities associated with over-the-air transmission, but the area of exposure is much greater due to transmission throughout the pager system.
- If reliability of pager message delivery is required, use at least a 1.5-way pager that gives a message acknowledge to receipt of message. The one way pager has no way to report message receipt.
- Digital pagers are somewhat less susceptible to attack than analog systems, but both are vulnerable to interception.
- Use the briefest message format possible. In terms of content, a numeric pager that requires a call-back is preferable to sending full messages on an alphanumeric system if the messages are not encrypted.
- Use of a standard wired telephone is preferable to the use of the Internet or a cellular phone for delivering messages to the service provider.
- At least one service provider (a team of SkyTel and V-One) provides an encryption service for over-the-air transmissions. The solution is better than no over-the-air security, but some exposure still exists within the service provider network and Internet connections.

### Desired Security Solution

- DII and certain NII customers require a higher degree of security in their pager network than is currently available. Sensitive information transmitted across a pager network should be encrypted on an end-to-end basis. This will require encryption capabilities at user terminals, i.e., the pagers. Reduced security involving over-the-air security only for

message content and addressing will be suitable for privacy applications on a case-by-case basis.

- Authentication of sending party and acknowledgment of receipt are desired characteristics.

## Best Commercially Available Solution

- Vendor solutions exist for provision of privacy level encryption using more advanced programmable user paging devices thus establishing a VPN environment for pager customers. However, the messages must be decrypted within the service provider network for routing purposes.
- If guaranteed delivery (or at least verification of delivery when it occurs) is a requirement, then a service provider must be selected that provides capabilities beyond the basic one-way paging systems.
- The recently announced provision of an elliptic curve public key cryptography key delivery system may assist in reducing the bandwidth overhead associated with Key Management Infrastructure (KMI) functions.

## Current Technology Gaps

- End-to-End encryption capability with minimal overhead encoding schemes.
- Short form rekey and SMI technology for authentication and key distribution.

### 5.2.5 Wireless Local Loop (WLL)/Wireless Public Branch Exchange (WPBX)/Cordless Telephones

Section 5.2.1 of this Framework discussed a wireless telephone environment where a user with a hand-held telephone roams throughout a cell structure controlled by a cellular service provider. This section describes a similar environment, but on a much smaller scale using what could be called a micro-cell or enclave structure. This section on wireless telephony defines a set of technologies and services that connect users to the wired circuit switched telephone network using local low power RF paths.

The three technologies in this section have been grouped together because of the similarities in their target environment, use of technology, and protocols. WLLs can provide telephone service to remote areas where a wired infrastructure does not exist, or can serve for reconstitution of communications when the wired infrastructure is damaged. Future deployment scenarios for the DoD foresee the use of wireless PBXs and cordless telephone equipment in remote areas or in tactical situations. The environment and range for the wireless PBX case is very similar to those for the WLAN.

A WLL can be described as a wireless replacement for the connection between the Central Office (CO) and user switching equipment. WLLs are often used to provide telephone service to

areas where laying cable is not practical because of terrain, or in remote areas where a microwave link or wireless modem is faster and easier to set up than a wired link to the CO. A typical configuration provides micro-cell concentrators within the local WLL service area with the RF links described above providing CO connection.

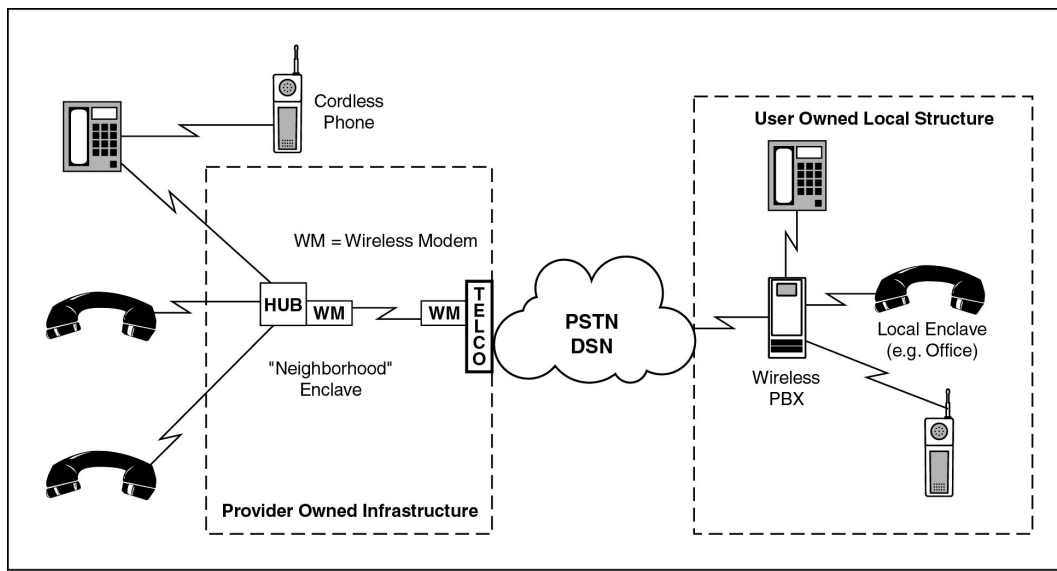
Wireless PBXs are often used in offices or manufacturing plants where individuals require mobility. A wireless PBX sets up a micro-cell structure where individuals carry a portable handset with them whenever they are away from their desk. Incoming calls are routed by the PBX first to their desktop phone, then to their portable phone. In essence, the portable phone is just an extension of the desktop phone that can be used from anywhere in the site within micro-cell range. This setup is used frequently in applications like hospitals and large manufacturing plants. The ability to handle high user densities is what distinguishes a wireless PBX cell structure from a cellular phone system described in Section 5.2.1 (Cellular Telephone).

Cordless phones are the most common of these three devices, used primarily in a household or neighborhood environment. Unlike the handset used with a wireless PBX, a cordless phone is used simply as a replacement for the standard desktop telephone. Each base station interacts with a single handset. The phones also have very limited range, typically under 150 feet, but the range is expanding as new products are introduced.

### 5.2.5.1 Target Environment

Commercial application of the WLL is primarily envisioned for third world areas or remote locations where a wired infrastructure does not exist. In Government applications, a wireless PBX could be used by military personnel as a field tactical telephone system that does not require stringing of wires, or even as replacement for elements of the TRI-TAC system. Both WLL and wireless PBX systems can help forces restore sufficient telephone service to stay connected to a main operating base in the event of loss of wired communications capability as long as the forces and the main operating base are in relatively close proximity or within line-of-sight using wireless modem interconnection. Many other applications exist within the standard office environment for DII and NII customers, especially where other data networks interface with the wireless system in use. Security requirements in these systems vary based on the threat in the local area. Sensitivity of communications, the need for reliability, and the amount of controlled space around an area using a wireless PBX or a cordless phone will help determine the specific threat to the user. A WLL provides for RF connections over a much larger physical area than the wireless PBX or cordless phone.

Figure 5.2-7 shows an example of how a wireless PBX and WLL could be deployed to provide telephone access in different situations. The WLL case uses a service provider system infrastructure while the wireless PBX has a user owned system infrastructure (again similar to the WLAN).



**Figure 5.2-7 Wireless Telephony Environments**

## 5.2.5.2 Consolidated Requirements

### 5.2.5.2.1 Functional Requirements

#### Users/User equipment (PBX and Cordless)

- Users must be able to make and receive dialed calls within the range of the system.
- Users must be provided with the standard features of wired telephony.
- Reliability and availability of service should be no worse than for wired system.
- Users and handsets must have assigned ID numbers.
- Handsets must be portable.
- Security of both control channel and user information channel information must be assured. The link between handset and base station must be at least as secure as the traditional wired telephone link.
- Confidentiality of user information on the “talk” channel is required.
- Confidentiality of keypad information should be provided. This function would secure credit card transactions, PIN numbers, and other account numbers that are entered on telephone keypads.
- Confidentiality of signaling and call set-up information is desired.



## 5.2.5.2.2 Networking Environments

Converge mobile and fixed wireless capabilities into one flexible hybrid network.

## 5.2.5.2.3 Interoperability Requirements

Wireless PBX and cordless telephone handsets should ideally be compatible with cellular telephone infrastructure.

## 5.2.5.2.4 Anticipated Future Requirements

- In addition to telephone services, WLL will also be used to provide Internet and Intranet access to distant locations at Integrated Services Digital Network (ISDN) data rates as a minimum.
- Militarized versions of commercial systems that provide end-to-end Type 1 confidentiality, call authentication, and jam resistance.

## 5.2.5.3 Potential Attacks

### 5.2.5.3.1 Passive

- WLL signals will typically traverse long distances on the reach back to the wired infrastructure using microwave or wireless modem systems. The signals pass across potentially hostile areas, providing easy access for an adversary.
- Wireless PBX and cordless communications have similar vulnerabilities to those discussed in the section on cellular communications. Both voice and control channel information is vulnerable to intercept, although the intercept range is smaller.

### 5.2.5.3.2 Active

- System administration for WLL and wireless PBX is typically done on a PC at the user location. System administrator functions can also be performed from remote locations through an Internet or dial-in connection. In this situation, all administrator functions are vulnerable to attack from any network around the globe. Therefore, sufficient protections must be in place to prevent unauthorized individuals from accessing the system.
- Denial-of-service attacks through electronic jamming, while easily detectable with the proper monitoring equipment, can have disastrous effects in emergency or battlefield situations.
- Spoofing attacks through changes in dialing or transmission of false messages are possible.

### 5.2.5.3.3 Insider

- Modify cordless handsets.
- Change user privileges in system administration database.
- Adjust output power control in micro-cells.

### 5.2.5.4 Potential Countermeasures

Several techniques are available to provide bulk encryption for WLL signals on the reachback (to the wired infrastructure) channels. Because of the high power and long distances covered with typical WLL installations, it is difficult to control where the signal radiates. Therefore, some method for encrypting this link is essential. Standard link encryption technologies (protocol independent) can serve the purpose.

For wireless PBX and cordless telephone channels, handsets and base stations can be equipped with a crypto token, or smart card device to provide security between the handset and the base station. At a minimum, some sort of data scrambling or spread spectrum modulation technique must be used to ensure the wireless link is at least as secure as a traditional wired telephone link. Spread spectrum techniques can also provide increased resistance to electronic jamming. Addition of a software or hardware token could be used to provide data confidentiality, and I&A required for more sensitive transmissions.

### 5.2.5.5 Technology Assessment

Several manufacturers provide WLL and wireless PBX solutions today that implement all the common telephony functions including call waiting, call forwarding, three-way calling, and voice mail. Most of these systems are designed for the office environment, and provide security features comparable to those found in cellular phone networks. Unlike cellular phone technology in the U.S., wireless PBX systems primarily use one signaling protocol, Digital Enhanced Cordless Telecommunications (DECT). DECT began as a cordless phone protocol, and is now used in the U.S. and Europe for both cordless phones and wireless PBXs. In addition to DECT, some cordless telephones use other signaling protocols like CT-1 and CT-2. The Personal Handdyphone System (PHS) is a protocol used primarily in Japan and other Asian markets.

WLL systems are still in the early stages of market deployment. As the number of products on the market increase, and users in the DII become aware of the benefits of WLL and wireless PBX systems in previously unwired urban environments, more frequent deployments of these systems will occur.

### 5.2.5.6 Usage Cases

Other sections of this Framework have addressed several cases involving connecting equipment at one classification level to equipment at the same, or different classification level across both trusted and untrusted networks. These cases are clearly an IATF issue, and also apply in the wireless domain. However, use of wireless equipment interfacing with a wired network

does not significantly change the cases that were previously discussed. In general, some level of communications security is recommended for any equipment where there is a connection to a potentially hostile or unknown environment. In the case of wireless communications, all transmissions can be thought of as connecting to an unknown environment because of the nature of RF transmissions and the ease of signal intercept. Thus, the descriptions of each of the specific cases addressed in this Framework remain unchanged for the wireless environment. Wireless telephony calls are treated herein as system high connections to their environment.

## 5.2.5.7 Framework Guidance

### Desired Security Solution

- As a minimum for NII and DII applications, the wireless equipment must provide data security equivalent to the security provided on a wired link. Basic analog or digital modulation of a voice signal without any data scrambling or spread spectrum modulation makes wireless transmissions easy targets for interception.
- For sensitive data, these wireless telephone systems must provide the capability to use appropriate encryption techniques for the level of information being transmitted. Implementation using hardware or software tokens for user handsets is a possible solution.

### Best Commercially Available Solution

As discussed in the section on cellular telephony, the best current solutions involve using a user carried installable token, e.g., akin to the SIM card with a cellular GSM or PCS phone to provide user I&A. Some cellular telephones provide wireless PBX and cordless telephone handset connectivity.

### Technology Gaps

- Other than the minimal privacy provided by digital transmission of voice signals over the air, very few currently available systems provide any degree of data confidentiality or data integrity. User tokens or SIM cards could help provide user authentication and data confidentiality for cordless telephones and wireless PBXs between the handset and the basestation.
- In such an obvious military application, the capability to provide ruggedized components and high-grade security is needed.

This page intentionally left blank.

## 5.3 System High Interconnections and Virtual Private Networks (VPNs)

Many new options opened in recent years for providing alternative security mechanisms for protecting of DoD information systems. Receiving justifiable attention are application layer mechanisms that offer end-system to end-system security services with stronger binding of the end user to applications than has been possible with simple password mechanisms. The problem has been that although the promise of application layer security has been very high, realization of all the benefits has been difficult. That difficulty arises from the fact that most computer platforms use operating systems that offer only minimal trust mechanisms if any at all. Since these untrusted operating systems control the computer platform resources, malicious elements of such operating systems could affect the invocation of the application layer trust mechanisms in ways that defeat the desired information assurance outcome. Moreover, the platform responds to network port operations in software processes outside the control of the higher layer security mechanisms, leaving the platform open to network attacks.

The response to this lack of strong invocation and lack of protection of the network port is that invocation of security mechanisms must be checked outside the end system. Furthermore, this checker must be the gatekeeper for whatever is allowed to pass to the end system. This gatekeeper has recently taken the form of an application layer guard that implements firewall mechanisms while performing an invocation check on all information allowed outside the protected enclave. This guard, while effective for non-real-time applications on networks with low sensitivity, has been difficult to scale to highly classified networks and real-time mechanisms. This difficulty, along with growth in the use of commercial networks by private industry, has created a renewed interest in efficiently using security mechanisms to create an effectively private network across a public backbone. This is not a new strategy for DoD. However, the renewed vigor in the pursuit of such solutions is recent. This section outlines the options available for implementing virtual private networks and gives sufficient information to trade off the options.

Before the wide dissemination of Internet technology, networking between separate parts of an organization required a privately owned system of communications lines or leased fixed telecommunications services connecting the various entities. The number of techniques for providing communications between facilities has increased dramatically. While leasing telecommunications lines is still an option for those with specialized communications environments, there are many more cost-effective options. All major telecommunications vendors offer an on-demand virtual network service based on narrowband Integrated Services Digital Network (ISDN), frame relay, or Switched Multi-megabit Data Service (SMDS). Some vendors offer higher data rate services based on Asynchronous Transfer Mode (ATM) technology. Some organizations are using connections over the Internet. With all of these communication methods comes some risk of exposing private information to outsiders. Each method offers varying degrees of risk and differing amounts of protection used to mitigate the risks. The purpose of this section is to explore the possibilities and to offer guidance on to how information should be protected in transit across these networks.

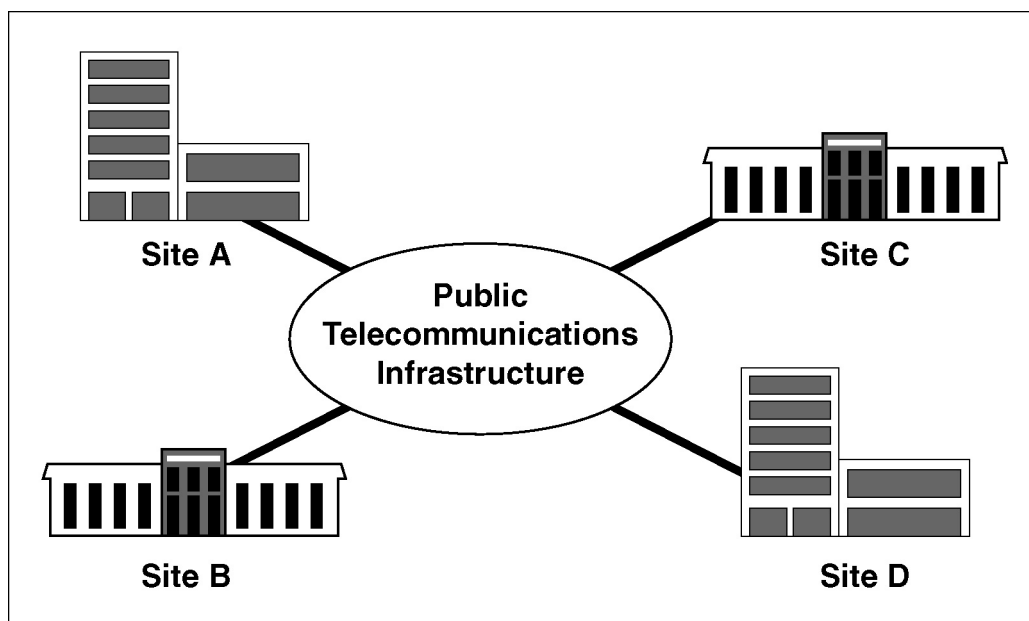
Some overlap is expected between the options presented here and in other portions of the IATF. This is particularly true for remote access of classified networks by lone users, and for High-to-Low interconnect. This overlap occurs because the various forms of networking discussed here are not unique. The particular end achieved is the result of a particular implementation of the underlying techniques.

One note on terminology. Throughout this section, the term “Type 1” strength cryptography is used. Traditionally this has meant government-developed or sponsored equipment containing security mechanisms that meet some minimum strength of implementation used where enough assurance mechanisms were in place to eliminate compromising failures. In the context that it is used here, it is generalized to include equipment from any source, provided that robust minimums of cryptographic strength and assurance mechanisms are included in the design. The exact definition of what these assurances and strengths must be is beyond the scope of this document.

### 5.3.1 Target Environment

A VPN allows the use of a public communications infrastructure in such a manner to exclude all entities outside a defined community. The communications may consist of leased lines, dial-up service, packet and cell switched connection-oriented networks, and or routed connectionless networks.

Figure 5.3-1 is deliberately vague about the type of communication infrastructure being used because a variety of infrastructures are possible.



**Figure 5.3-1 Target Environment Communications Infrastructure**

For example, the following infrastructures are among those available today:

- If the service is switched and connection oriented, it can be frame relay or ATM

## UNCLASSIFIED

System High Interconnections and Virtual Private Networks (VPNs)  
IATF Release 3.0—September 2000

- If it is dial-up service, it can be based on ISDN or Digital Subscriber Line (DSL)
- If it is packet-switched and connectionless, it can be Internet or SMDS
- If the service is leased line, it can be Digital Service, Level Zero (DS-0), DS-1, Fractional DS-1, Burstable T-1, DS-3, Synchronous Service Transport, Level Three (SST-3), or higher rates in North America. Table 5.3-1 provides additional information for each these.

**Table 5.3-1 Digital Service Standards**

Digital Standards	Definition
DS-0	In the digital hierarchy, this signaling standard defines a transmission speed of 64 Kbps. This is the worldwide standard speed for digitizing one voice conversation; (i.e., converting one analog voice channel into a digital signal. It is derived from using pulse code modulation (PCM) and sampling the voice channel 8,000 times a second. This signal is then encoded using an 8-bit code. Thus, 64,000 bps is derived from 8-bits times 8,000 times per second.
DS-1	In the digital hierarchy, this signaling standard defines a transmission speed of 1.544 Mbps. A DS-1 signal is composed of 24 DS-0 channels. DS-1 is often used interchangeably with T-1, which is the U.S. equivalent of E-1. T-1 is a Bell system term for a digital carrier facility used for transmission of data through the telephone hierarchy at a transmission of 1.544 Mbps. E-1 is the European equivalent of a T-1 circuit. E-1 is a term for digital facility used for transmitting data over a telephone network at 2.048 Mbps.
Fractional DS-1	A DS-1 circuit in which a fraction of the 24 DS-0 channels are used; (i.e., between 64 Kbps and 1.536 Kbps. If a full DS-1 circuit is 24 DS-0 channels at 1.544 Mbps, a 1/8 Fractional DS-1 is four DS-0 channels at 256 Kbps, a 1/2 Fractional DS-1 is 12 DS-0 channels at 768 Kbps and 2/3 Fractional DS-1 is 16 DS-0 channels at 1,024 Kbps.
Burstable T1	This service is a billing scheme. It is an unshared, non-fractional T-1 line running at 1.544 Mbps. While a DS-1/T-1 customer has the full capacity of the line (24 DS-0 channels at 1.544 Mbps) any time he desires, the customer is billed only an average usage computed from periodic samplings of the input and output data rates on the link.
DS-3	In the digital hierarchy, this signaling standard defines a transmission speed of 44,736 Mbps. A DS-3 signal is composed of 673 DS-0 channels. DS-3 is often used interchangeably with T-3, which is the U.S. equivalent of E-3. T-3 is a Bell system term for a digital carrier facility used for transmission of data through the telephone hierarchy at a transmission rate of 45 Mbps. E-3 is the European equivalent of a T-3 circuit. E-3 is a term for a digital facility used for transmitting data over a telephone network at 34 Mbps. Also available is a Fractional DS-3 service in which a fraction of the 28 DS-1 channels are used; i.e., (between 1.544 Mbps and 43,232 Mbps. Other digital service levels are available; (e.g., DS-2. 96 DS-0 channels at 6,312 Mbps; DS-4, 4,032 DS-0 channels at 274,760 Mbps.
SST	This is a SONET-based, private line transport product that offers high-capacity channels for synchronous transmission at transport line rate from 155.52 Mbps to 2,488 Gbps. It enables the interfacing of asynchronous networks with synchronous networks.

## UNCLASSIFIED

System High Interconnections and Virtual Private Networks (VPNs)

IATF Release 3.0—September 2000

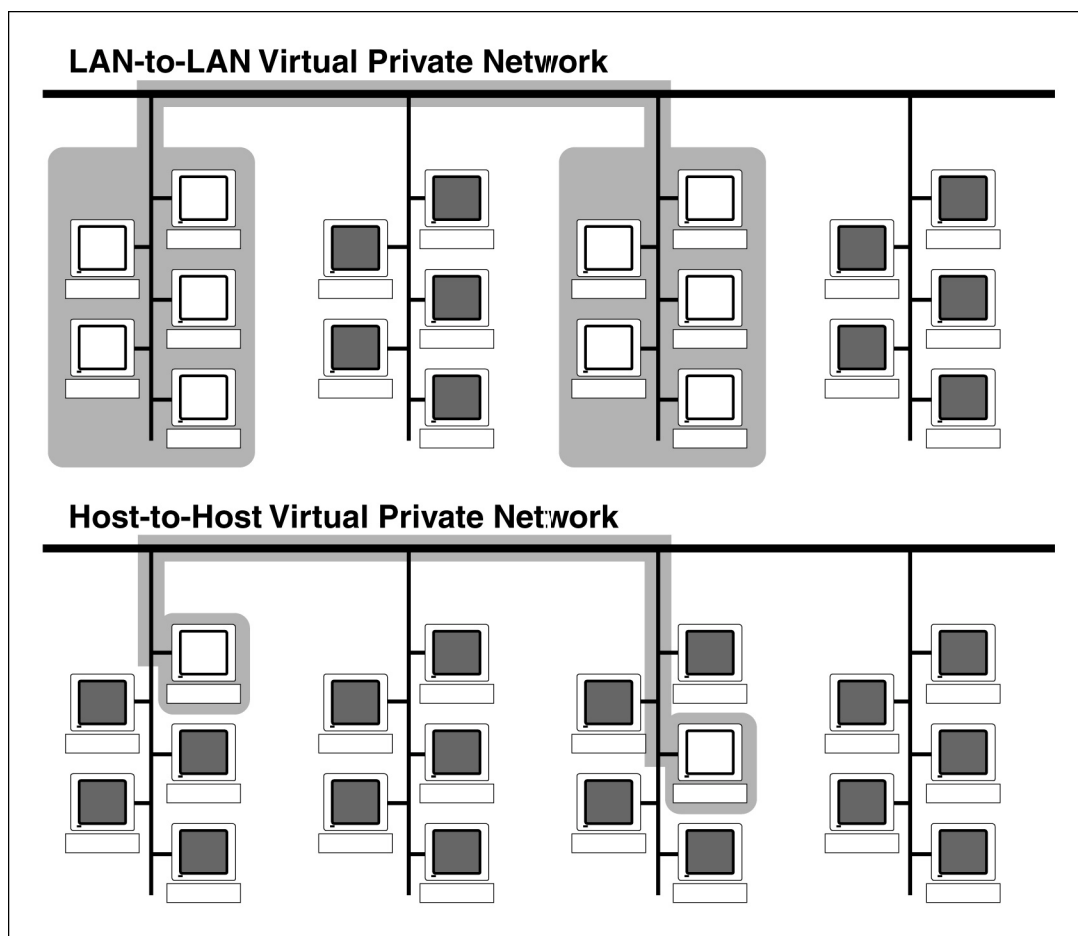
Digital Standards	Definition
DSL	<p>DSLs are point-to-point public network access technologies that allow multiple forms of data, voice, and video to be carried over twisted-pair copper wire on the local loop between a network service provider's central office and the customer site. Included are Asymmetric Digital Subscriber Line (ADSL), Rate-Adaptive Digital Subscriber Line (R-ADSL), High Bit-Rate Digital Subscriber Line (HDSL), Single-Line Digital Subscriber Line (SDLS), and Very High Bit-Rate Digital Subscriber Line (VDSL). Collectively, the DSL technologies often are referred to as xDSL. ADSL is an xDSL technology that allows more bandwidth downstream—from a network service provider's central office to the customer site—than upstream from the subscriber to the central office. ADSL is ideal for Internet/intranet surfing, video-on-demand, and remote LAN accesses. R-ADSL is an xDSL technology that adjusts dynamically to varying lengths and qualities of twisted-pair local access lines. R-ADSL makes it possible to connect over different lines at varying speeds. HDSL is an xDSL technology that is symmetric, providing the same amount of bandwidth both upstream and downstream. Due to its speed—1.544 Mbps over two copper pairs and 2.048 Mbps over three copper pairs—TELCOs commonly deploy HDSL as an alternative to repeated T-1/E-1 lines. SDLS is an xDSL technology that provides the subscriber only one DSL line. VDSL is the fastest xDSL technology, supporting a downstream rate of 13 to 52 Mbps and an upstream rate of 1.5 to 2.3 Mbps over a single copper-pair wire. Maximum operating distance for this asymmetric technology is 1,000 to 4,500 feet. The VDSL bandwidth could potentially enable network service providers to deliver high-definition television signals in the future.</p> <p>Note: TELCO is a generic term for local telephone company operations in a given area.</p>

No matter what the underlying communications scheme, the desired result is to connect separate pieces of a larger organization in a manner that provides unimpeded communications between the pieces of the organization, denies access to the information within the pieces by any outside organization, and provides for the privacy of information as it traverses the public infrastructure.

Many people make the assumption that a VPN is a distributed enterprise network connected across a public Internet but separated from that Internet by an encrypting firewall. This use of the term precedes the definition of Internet Protocol Security (IPSec) that is the basis of the present generation of encrypting firewalls. The three major telecommunications carriers offer a virtual private networking service that combines voice and data features, billing, access, screening, and re-routing capabilities but does not have any inherent encryption mechanism.[3] This chapter uses a broader definition of VPN that encompasses any means of using public communications infrastructure to manifest an apparently private network.

In the context of this IATF, there is little difference between a system-high interconnect and a VPN. Possibly the only real difference is that the end systems have implemented a private network with an wholly owned infrastructure or the end systems are use a shared backbone based on some publicly offered service. Although some state that use of a provisioned service like DS-3 or Synchronous Optical Network (SONET) is a system-high interconnect, these services are multiplexed onto a public backbone, managed by a public entity, and the routes can slowly change in response to some network conditions. Therefore, even this type of networking represents the creation of a VPN across a public switched backbone.





**Figure 5.3-2 Local Virtual Private Network Architectures**

## 5.3.2 Consolidated Requirements

The present requirements are derived from operating scenarios of present system-high networks based on use of leased line services and on an interconnect model that uses the Internet. Anticipated requirements are derived from plans for the far-term Defense Information Systems Network (DISN), technology developments from Defense Advanced Research Projects Agency (DARPA) and the Global Grid community, plans stated by telecommunications vendors, and aggressive Research and Development (R&D) networks such as those pursued under the Nuclear Stewardship program.

### 5.3.2.1 Functional Requirements

Near-term functional requirements are as follows:

- Must support connection of separated entities across public infrastructures (site-to-site model) or within private facilities ((Local Area Network ) LAN-to-LAN or host-to-host model).

- Must support classified operations or unclassified operations.
- Must support standards-based network operations.
- Must keep network information confidential and integral while in transit.
- Must prevent entities outside the private facilities from gaining access to those facilities.
- Must use techniques that support scaleable communications rates from kilobit per second rates to OC-192 (10Gbps) and beyond.
- Must transport is primarily data including voice, video, imagery, and data.
- Must optionally provide data integrity.

Mid to far-term functional requirements are as follows:

- Must support quality of service in the telecommunications must be supported.
- Must support data rates for specialized applications that exceed 13 Gbs by the middle of the next decade.
- Must support information that is mixed voice, video, and data.
- Must connect non-uniform security policies. As more risk management philosophies are developed for administering security within network domains, security policies can be expected to diversify even within similar classification level of networks. These discrepancies will result in additional security requirements on VPN architectures.)

## 5.3.2.2 Networking Environments

This section provides as a local reference regarding environments in the context of the virtual private networking arena.

Two networking environments are currently dominant. The first is link layer connection over leased lines and the second is Internet Protocol (IP) packet routing over the Internet or ATM wide area networks. Although frame relay and SMDS technologies have made significant inroads into the business community, they have been used rarely for classified communications. This has been attributed in part, to the lack of native mode security systems for these means of communication but also because there have been alternative means of achieving security services could be used without affecting the functionality of the network.

Networking environments will undergo drastic changes over the next few years. With this revolution will come an explosion in the number of networking technologies. Although the IP and provision networks of today will not disappear, they will be joined by newer technologies and by variations of the old technologies. The present IP version 4 will evolve to incorporate bandwidth reservation schemes in an attempt to add quality of service attributes to deliver business-quality voice and video applications over the Internet. Other users will move to ATM networks because they are designed to deliver quality of service for these same applications. A war for market share will ensue between these networking technologies. The outcome of this battle is not clear. Currently, neither technology fully achieves all of its promises. The expected result will likely be a coexistence of these technologies.

As wireless network technologies evolve, there is likely to be a specialization of IP for the mobile environment that will require some level of gateway to the wired portion of the Internet.

Speeds of connectivity will increase. The maximum available today in a standardized format is an STS-48/STM-16 signal at 2.5 Gb/s and some initial deployment of an STS-192/STM-48 signal at 10 Gb/s. These signals will be wavelength division multiplexed up to 40 and 80 Gb/s. The affordability of such large bandwidths is certainly a major issue. However, a few programs have identified communications requirements of greater than 10 Gb/s. The most easily referenced example is Department of Energy's (DOE) Nuclear Stewardship Program. To support simulations of aging effects in stockpiled nuclear weapons, it is estimated that computational capacities of 0.1 Petaflops are required, backed by 13 Gb/s communications between the DOE weapons laboratories.

### 5.3.2.3 Interoperability

A trend within the DoD is to break down barriers to connectivity rather than put more barriers in place. As a result, the natural segregation that would occur between entities in different communications environments, between entities communicating at different rates, and between those entities using different networking architectures is breaking down. Therefore, one must assume that a secure means of exchanging information between the various networking architectures is required.

Another interoperability issue is DoD trend toward breaking down barriers between networks operating at different levels of classification and assurance. Although, this is a multilevel security problem and not a virtual private networking issue, the solutions must be mutually supportive.

### 5.3.3 Potential Attacks

The attacks listed here are those primarily of concern to systems protected at network layers and below. One interesting paper, although written primarily about a particular implementation of IP-based security, presents an open tutorial of many issues must be considered when implementing network layer security solutions.[1] Although, the author often assumes that an adversary already has access to a private resource and therefore presents a pessimistic picture, the subject matter at least considers many security issues that are often ignored. This paper is used as a reference throughout this section.

Attacks against networks vary greatly regarding the techniques and results. While some try only to uncover private information, others try to disrupt operations, disseminate misinformation, and gain access to resources.

#### 5.3.3.1 Passive Attacks

The primary concern with passive intercept attacks is the loss of information confidentiality while in transit across the network. Basic privacy rules to prevent inadvertent disclosure are insufficient for DoD. Recent reports show that cryptanalytic capability is available in the public domain as witnessed by the June 1997 collaborative breaking of 56-bit strength Data Encryption Standard (DES). Although, the near-term threat to large volumes of traffic is questionable given the number of machines and hours involved, it does show the vulnerability of any single

transaction. Therefore confidentiality mechanisms must pass some measure of minimum strength to be acceptable. However, that is not the only concern. Some military operations require the element of surprise. Therefore, one must assess the possibility of passive observation of network operations giving indications and warnings of impending actions. Such indications may be who the end parties are in an information exchange, a change in the volume of traffic or traffic patterns, or the timing of information exchanges in relationship to external events. The resulting potential security requirements are strong confidentiality and traffic flow security.

### 5.3.3.2 Active Attacks

This class of attacks may involve end systems or infrastructure. The most obvious network-based attack is the attempted login to a private computational resource. Bellovin shows how the ability to splice messages together can be used to change information in transit and cause desired results.[1] In the financial community, it could be disastrous if electronic transactions could be modified to change the amount of the transaction or redirect the transaction into another account. Re-insertion of previous messages could delay timely actions. Bellovin also brings up the issue of chosen plain text attacks<sup>1</sup> that can be used to bypass encryption mechanisms. [1]

Denial-of-service attacks can be minimized by choice of network technologies. Any network that supports dial-up connections or routing of information can be used to deny service by flooding an end point with spurious calls or packets. More sophisticated attacks can involve manipulation of network elements.

The following are resulting potential countermeasures.

- Strong access control
- Continuous authentication
- Integrity of information
- Replay prevention
- Network availability

### 5.3.3.3 Insider Attacks

Many insider attacks are possible in a VPN. This is an architecture that concentrates on control of outside access. There is no additional mechanism to inhibit a person with legitimate access to a system from accessing more private areas of the VPN. A malicious insider could use covert channels to signal private information outside the VPN. However, there are many other avenues for a malicious insider to wreak havoc with an information system. Another threat that must be considered is the introduction of malicious code into a protected enclave. Such code can be easily imported through shrink-wrapped untrusted software, users swapping media with machines outside the enclave, or other paths that are implemented to import information

---

<sup>1</sup> Many attacks are aided by making a machine encrypt plaintext chosen by the attacker. Many cryptanalytic attacks depend on the attacker being able to choose the plaintext to be encrypted [2]

from outside the VPN. Although many precautionary security requirements could be taken that are outside the scope of the virtual private networking scenario, the resulting potential security requirements for the VPN are establishment of security domains within the VPN, and control of covert channels.

## 5.3.4 Potential Countermeasures

Privacy is maintained by appropriate use of confidentiality mechanisms. While application layer mechanisms can provide information confidentiality for classified and other critical applications, the problem with assured invocation of these mechanisms makes it difficult for these mechanisms to provide primary confidentiality mechanisms. The strength of confidentiality mechanisms for classified applications must be sufficient to withstand national laboratory strength attacks.

If traffic flow security is required, the best mechanism is one that prevents all insight into changes in traffic patterns. Therefore, the best mechanisms are link layer mechanisms on constant bit rate leased lines. Alternatively, lesser degrees of traffic flow security can be afforded by aggregating traffic through secure tunnels and by using traffic shaping mechanisms.

Many network attacks that involve manipulating cipher text or splicing information units can be countered by strong data integrity mechanisms and continuous authentication of the data channel. Replay can be prevented with cryptographic mechanisms that use timestamps or incrementing of counters to limit the acceptability of prior messages in the end systems. Continuously authenticated channels can prevent insertion of information into the channel that such insertions could permit short plaintext attacks that would allow cryptanalysis by guessing known responses to known short messages.

Prevention of denial-of-service attacks is often in the hands of the network provider. Use of provisioned networks will prevent many denial-of-service attacks because the general population is unfamiliar with the management mechanisms in networks. However, there is little in present infrastructures to prevent manipulation of network hardware. The router authentication being implemented in the DISN is a start toward decreasing the vulnerability of networks to manipulation of network management information. Similar moves are being proposed within the Security Working Group of the ATM Forum for control of ATM switch configuration messages. Neither of these techniques is widespread so the network remains vulnerable to hacking.

Virtual private networking architectures provide little protection against the insider threat. Malicious insiders or malicious code introduced into the network all operate above network layers. These threats must be handled by higher layer services. If insider threats are a concern, the security implementation should also consider inclusion of firewalls, end-system-based privacy mechanisms, and protection mechanisms over the wide area network that limit exposure to covert channels.

## 5.3.5 Technology Assessment

There are many ways to implement a secure VPN. The easiest method for categorizing the options is to look at the possibilities as one moves up the protocol stack in a network. For purposes of this IATF, the discussion starts at link layer protocols where framing can take place. This is the lowest layer that can be transported through a standardized public infrastructure. The discussion stops at the transport layers. It should be noted that transport layer security services normally could only exist in end systems unless, at some future point, a transport layer proxy is created in a gateway device.

### 5.3.5.1 Layer 2 Protected Networks

The option of protecting a network at layer 2 is possible only if the owner has installed or leased a dedicated communications facility between sites. The security services that one achieves with a layer 2 protected network are strong site-to-site authentication, confidentiality, and a continuously authenticated channel. In most cases, one also achieves traffic flow security. An optional security service may be some data integrity functions or at least an anti-spoof capability.

A layer 2 protected network, given present protocol suites, cannot provide any true end-user authentication. It cannot provide any degree of privacy between users within the protected network at a reasonable expense. All switching and routing facilities will be Red facilities unless supplemented by other security mechanisms. This option contains no provisions for limiting information flow between facilities. If a firewall or equivalent function is required, it inserted before the link encryption mechanisms.

Given the limitations outlined above, layer 2 protection for networks could easily be dismissed as not useful. However, some security mechanisms can't not easily be used in higher layers. The first mechanism is traffic flow security. If a user is concerned about receiving indications and warnings about impending actions, traffic flow security is imperative. Although, some traffic flow security is possible using rate shaping of information, this technique requires non-standard applications and protocol stacks, which could entail significant life cycle costs.

The second mechanism not available in higher layer is the limitation in the number of covert channels. Covert channels are often viewed as either the gravest of threats to our information systems or a hobgoblin to be dismissed with a wave of the hand. The reality is that accreditors must have to evaluate the threat of covert channels to their particular information system and determine the desired level of protection against the threat. Although, a detailed discussion of any of these vulnerabilities is outside the scope of this paper, it does not take too active an imagination to postulate the existence of covert channels given that any field in an packet that can be modified or any parameter of transmission that can be varied is a potential covert channel. A layer 2 protected network removes all covert channel classes encompassing length of information transfer, timing of information transfers, and addressing of information transfers. Remaining covert channels can arise from the ability to exploit incompletely defined transport overhead and will be stemmed by the ability to control access to the overhead.

Another desirable property is that the simplicity of the design of link layer systems means that it is easier to achieve a target throughput at the link layer than at any other layer. As users

reach for the limits of available communications technologies, it is more likely that a link layer solution will be the most acceptable solution. Table 5.3-2 summarizes the positive and negative characteristics of layer 2 protected networks.

**Table 5.3-2 Characteristics of Layer 2 Protected Networks**

Positive Characteristics	Negative Characteristics
Highest speeds possible	Highest communications costs
Highest protection against traffic analysis	No protection against cascading of networks
Highest protection against covert channels	No protection against insiders
Fewest avenues for network-based attacks	Can only authenticate from site to site
Continuous site-to-site authentication	Requires carrier to reconfigure network to add new nodes

- 1) *SONET*—SONET is the standard in the United States (U.S.) for trunking of data at rates greater than 45 megabits per second. It is delivered in multiples of 51.84 Mb/s with the minimum multiple being three. This service is referred as a (STS-3.) If the entire capacity is treated as a single data container, the service is referred to as STS-3c, where the c denotes a concatenated service. The international version of this service is Synchronous Digital Hierarchy. The basic unit of service is a Synchronous Transport Multiplex, which is the equivalent of the SONET STS-3c transport. Present widespread deployment supports 155, 622, and 2488 Mb/s transmission rates. Initial deployments of SONET at 9952 Mb/s have occurred. Approximately 3.33 percent of the data flow is devoted to transport overhead. Another 1.11 percent is devoted to path overhead in non-concatenated channels.

Presently, only government-developed equipment is available to secure SONET networks. SONET Key Generators encrypt the data payload providing for strong confidentiality and complete traffic flow confidentiality. Data integrity must to be handled at higher layers. SONET overhead passes through the system unaltered or, alternatively, only minimum fields are passed through the system undefined and network control channels are cleared. The operators of local SONET networks decide the level of transport overhead flow between local and wide area environments. A commercial device has been developed to meter these interactions between local and wide area SONET network but the future of the device is not certain. No known commercial SONET encryptors exist at this time. However, a commercial entity has expressed interest in providing services based on such a device.

- 2) *Sub-SONET Rate Services*—The widespread data trunks in the U.S. are fractional DS-1, DS-1 at 1.544 Mb/s, and DS-3 at ~ 45 Mb/s. These services represent a multiplexed hierarchy for combining 64 kb/s voice channels into higher order trunks and eventually into SONETs adapted to direct transport of non-voice data. The transport overhead varies from 1.4 percent for DS-1 service to 3.9 percent for DS-3. Trunk services are protected by a series of standard government-developed encryption equipment. These encryptors have been the basis of numerous VPNs based on provisioned services. In addition, numerous commercial offerings have seen a limited success in the market place. Commercial link encryptors are ripe for evaluation for possible use in layer 2-

protected VPNs. Similar to the SONET devices described above, such link encryptors provide strong confidentiality, continuously authenticated channels, and traffic flow protection. They may also provide data integrity based on error extension properties of the encryption mechanism.

An interesting alternative to securing constant provisioned services is to apply an ATM-based solution. Because ATM can transport constant bit rate services, it is possible to use a cell-encryption-based technology to provide encryption services for link layer protocols. Many technical issues must be considered in the actual implementation of this technique. Among others, how the physical link is manifested at the service access point and relative costs are important considerations. Such a solution may not have all the security properties of traditional link encryptors. A discussion of the security properties of ATM will be included in a later release of this document.

- 3) *N-ISDN*—Narrowband Integrated Services Digital Network (N-ISDN) is a digital data transport system. It can be supplied in several forms including basic rate and primary rate services. Basic rate service consists of two data channels and one signaling channel with a combined capacity of 144 kilobits per second (kb/s). In the United States primary rate service consists of 23 data channels and 1 signaling channel for a total capacity of 1.544 Mb/s. Europe and Japan, use a different standard for primary rate service. Government equipment is being designed for N-ISDN. This device was initially prototyped as a single data channel and a single signaling channel and has since been followed with a version with two data channels, and one signaling channels. No known commercial devices exist for native N-ISDN security. Security services available for N-ISDN depend on how security is invoked. Security can be implemented by encrypting complete data channels. Such an implementation would have security properties similar to the link encryption devices discussed above. N-ISDN can also be used for multiplexed data transport. In fact, this transport is the basis of the commercially successful frame relay service offered by many carriers. If security is invoked at this layer, security properties will be the same as those discussed in the layer 3 section to follow.

N-ISDN is used as a low bandwidth connection between end systems and as a medium speed dial-up temporary connection between fixed and mobile systems. Direct dial-up secure N-ISDN represents a reasonable protection for dial-up access into a secure enclave, provided that policy allows such connections, strong user authentication is invoked, and procedures are put in place to protect classified information on a remote system while outside a protected enclave.

- 4) *Analog Phone Service for Data Transport*—Analog phone service requires a digital modem for transport of information across the analog link and is available as a dial-up medium for low bandwidth temporary connections. Newer modem technologies represent nearly the same capacity as an N-ISDN data channel without the set up charges and communications cost associated with N-ISDN. Commercial prototype encrypting modems have been developed for such secure data connection use and represent a reasonable method of providing a temporary link to a VPN, provided that strong user authentication is part of the connection process.

An alternative to the encrypting modem is the use of the data port of the government-developed secure telephones. Part of the authentication scheme for a government secure voice equipment is the voice recognition between speakers. A totally automated



system could bypass this important function. Many dial-up functions in low-cost computers accept manual dialing. A possible security policy would be to require audio identification of the sender before going secure or to require an augmenting strong authentication during log-in.

- 5) *Voice Transport*—Voice networks are often disregarded by the data network community, but in the DoD they still carry a large volume of secure traffic. Modern secure phones are based on digital representations of voice that are encrypted and sent across the network by digital modem. This is true whether the end system is connected to an analog service like Plain Old Telephone Service (POTS) and analog cellular service or a digital service like N-ISDN or newer digital cellular technologies. The distinction between voice networks and data networks is expected to diminish in the next few years. N-ISDN, ATM, digital cellular, and Internet phone are already blurring the lines. Government secure voice architectures have unified secure interoperability across most voice transport mechanisms. The exceptions to this rule are Internet Phone and native ATM voice transports. An area ripe for work is the extension of secure voice architectures into these newer network technologies.

### 5.3.5.2 Layer 3 Protection across Public Networks

Layer 3 networks support dynamic routing and switching of information. For the purposes of the IATF, this discussion primarily covers IP and ATM transport. For this reason, the discussion is not complete. Network protocols like Network Basic Input/Output System (NETBIOS) and Internet Packet eXchange (IPX) are not covered. In addition, ATM spans a range of network layers. If implemented as a permanent virtual circuit, it becomes a strict layer 2 entity. In many implementations, ATM is used below layer 3 but above the Media Access Controller becoming the equivalent of about a layer 2.5 entity. Prototype applications are capable of completely replacing layer 3 solutions. Because of the cell switched nature of ATM, it is closer in properties to the pure layer 3 solutions and is therefore handled in this section. A protection philosophy based on layer 3 type networks offers the end users more affordable communications costs than layer 2 protected systems. A layer 2-protected system requires the provisioning of a new communications line and the acquisition of a pair of protection devices enable the new connectivity. With a layer 3-protected system, one only has to enable the access control mechanisms to allow the new connectivity. This comes at a cost of a higher risk of vulnerability to traffic analysis and the exposure to covert channel problems and directed network-based attacks. Table 5.3-3 summarizes the characteristics of layer 3 protected networks.

## IP Network

Only one widespread Type 1 system provides layer 3 protection for networks—the Network Encryption System (NES). This system uses a security protocol called SP-3 to encapsulate and transmit information securely across the Internet. NES has its own unique IP address and a broadcast address. When information is encapsulated, the outer IP envelope contains only gateway-to-gateway addresses. Therefore, end system identity is not available in the public Internet.

**Table 5.3-3 Characteristics of Layer 3 Protected Networks**

Positive Characteristics	Negative Characteristics
<p>Some billing models charge by volume of traffic allowing greatest control of cost</p> <p>Most flexibility in adding new nodes to network</p> <p>Continuous site-to-site authentication possible</p>	<p>Traffic analysis easy under some configurations</p> <p>No protection against cascading of networks</p> <p>No protection against insiders</p> <p>Many covert channels for exploitation</p> <p>Many denial-of-service attacks possible under some implementations</p>

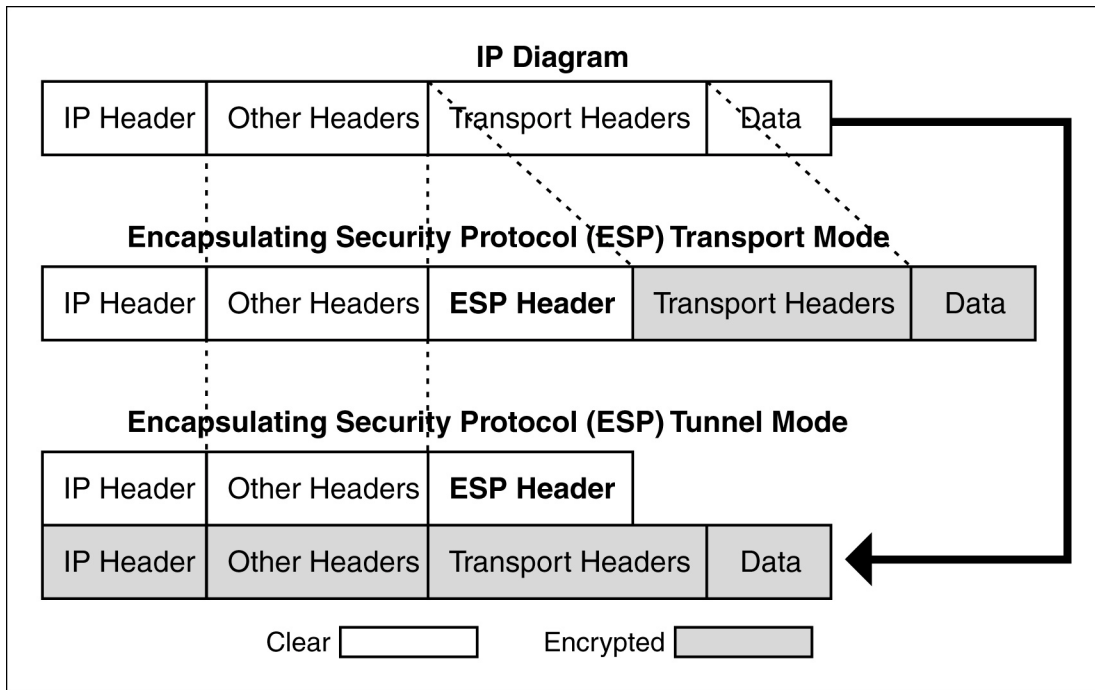
For this method to work, the device contains a configuration table that maps end system addresses to gateway addresses. The security services provided are site-to-site confidentiality, site-to-site authentication, and site-to-site integrity. Traffic flow protection of the aggregate data flow is not provided, although it is possible to write specialized applications whose purpose is to smooth the traffic flow across a site-to-site flow.

Numerous commercial IP encryptors also exist. These older commercial systems tend to have many proprietary features that preclude interoperability of equipment. Because of this lack of interoperability, it is not recommended that older commercial IP-based encryption systems be studied for securing DoD systems. For immediate applications requiring a layer 3-protection mechanism in support of the flow of classified information, NES is the only available solution.

There is a potential for a widespread IP layer encryption solutions based upon what has been called IPSec. IPSec is the security framework that has been standardized by the Internet Engineering Task Force as the primary network-layer protection mechanism. IPSec consists of two parts, an Authentication Header (AH), whose purpose is to bind the data content of IP frames to the identity of the originator, and an Encapsulating Security Protocol (ESP) for privacy. The AH is intended to be used when integrity of information is required but privacy is not. ESP is intended to be used where data confidentiality is required. The draft Request for Comments (RFC) that define IPSec architecture states that if data integrity and authentication are required with confidentiality, then an appropriate security transform should be used that provides all services. The minimum set of protection mechanisms consists of the DES for confidentiality and the hash algorithm MD-5 for authentication. The standard does provide room for negotiating alternative protection mechanisms through use of the Internet Key Exchange Protocol (IKE). IKE provides both a framework for creating security associations between endpoints on a network and a methodology to complete the key exchange. At least one published paper points out potential security concerns about using IPSec default security mechanisms. [1] The author points to occasions where the integrity functions of DES in Cipher Block Chaining mode can be circumvented with the right applications by splicing of packets. [1] The referenced paper recommends that AH and ESP be used together instead of individually.

ESP defines two methods of encapsulating information: tunnel mode and transport mode. Tunnel mode, when used at an enclave boundary, aggregates traffic flow from site to site and thereby hides end system identity. Transport mode leaves end system identity in the clear and is most advantageous when implemented at the end system. Figure 5.3-3 shows where the ESP header is placed within an IP datagram for IP version 6. In the more ubiquitous IP version 4, the section marked Other Headers does not exist. The AH precedes all non-changing end-to-

end headers. If one wanted to follow Bellare's suggestion and use AH with ESP, the authentication header must immediately precede the ESP header. [1]



**Figure 5.3-3 IP Layering Encryption Methods**

Although, no government-sponsored equipment currently implements IPsec, one such device is under development. TACLANE is an IPsec and ATM encryptor that will be certified to handle classified information. It uses the ESP tunnel mode without the AH. It also does not implement the default IPsec algorithms of DES and keyed MD-5. Because hard-wired security policy states that DES and MD-5 are not strong enough for Type 1 grade security. TACLANE always negotiates to higher-grade security mechanisms or does not commence data transmission. A follow-on development for the TACLANE program will provide fast Ethernet cards for TACLANE and increase its encrypted IP throughput to 100 Mbps.

It is recommended that all future IP security equipment should be IPsec compliant. The primary confidentiality mechanisms should be implemented in security gateways that support no user-level processes.

No Type 1 grade IPsec-compliant commercial encryptors currently exist. Even in current government developments, there are technology gaps for devices that can handle full Ethernet bandwidths, 100 Mb/s Ethernet bandwidths, and Gigabit Ethernet bandwidths. In the commercial arena, there are many IPsec implementations for individual end systems and those that incorporate into firewall capability. Both of these implementations will require Type 1 grade equipment.

## ATM

ATM security was developed in anticipation of requirements for high-quality multimedia communications. The flexibility of the transmission mechanism make it possible to tailor the

security features of the system depending on how ATM is used. The standardization process for security in ATM is not as well established as that for the IP community, although some basic features and cryptographic modes have been defined through the Security Working Group of the ATM Forum. One challenge that can be issued to the ATM community is to bring the definition of an “ATMSEC” to the same level of sophistication as that developed in the Internet Engineering Task Force.

Some of the main differences between ATM and IP include the following. ATM relies on a call set-up mechanism or explicit provisioning while IP routes are discovered en route. ATM relies on the state of a connection, while IP (especially version 4 IP) is stateless. ATM fixes cell size while IP uses variable size packets. IP frames carry end-to-end address information whereas ATM cells only carry local identifiers between each pair of switches. Quality of service in ATM is determined by availability along the entire route whereas IP quality of service is based solely on admission control to the network.

The primary motivations for considering ATM security solutions the need to integrate high quality voice, video, and data applications and the need for quick implementation. Although the abilities of ATM are more apparent at the high end of communications, the mechanism scales across a wide range of operating rates.

Because IP packets can be reordered in transmission, each packet must contain sufficient information to enable synchronization of security mechanisms. ATM security can rely on the state of the connection to maintain synchronization. If the implementation is aware of ATM adaptation layers, information is available to deal with a limited amount of cell loss while maintaining synchronization. IPSec defines per packet authentication schemes through the AH. ATM security, as defined to date, does not have the equivalent function. Anti-spoof functionality is available that relies on higher layers to complete authentication, but the degree of protection is not the same as IP using the AH.

Because ATM can be implemented in so many ways and because the security services differ for each implementation, the options are discussed individually.

ATM can be used in a Constant Bit Rate (CBR) mode to connect enclaves emulating layer 2 trunk traffic. When ATM is this way while configured as a Permanent Virtual Circuit (PVC), all of the security services of secure provisioned link communications are available but provide more flexibility for upgrading service as required. If Switched Virtual Circuit (SVC) service is available at the enclaves, then potential denial-of-service attacks must be handled. Enclave-to-enclave IP over secure ATM (RFC 1483) has the same security attributes as IPSec in tunnel mode. Site-to-site identification is possible but the identity of end systems is hidden within the tunnel. Traffic rate is visible to the outside world but aggregation of large amounts of traffic and traffic smoothing can help obscure traffic flow information. Because of this similarity, this section refers to such a mode as a tunneling mode of ATM despite the lack of a formal definition. End-system-to-end system secure ATM has security properties similar to IPSec transport mode. Complete end system identification is possible and individual traffic flows are discernible. Secure virtual paths allow, end system identity to be hidden within a secure signaling channel within the virtual path. Though individual traffic flows will be discernible on the wide area network, there will be no information to tie the flow to an originator within the enclave except for perhaps stimulated events. Similar to the tunneling case, when end to end-user information is available, this section refers to that ATM transport mode as a tunneling mode.

The splicing attacks that Bellovin attributes to IPSec encapsulating security payloads may also be possible with ATM Forum-recommended encryption mechanisms.[1] This is an area for further study. If such an attack is possible, there is no equivalent to the AH to counter the threat. It is important to note that even if such attacks are possible with the ATM Forum-recommended modes, that such attacks need not exist with all algorithm suites.

Government-sponsored equipment for securing ATM SVCs and PVCs are available for data rates up to 622 Mb/s. A Type 1 interim system was developed for a single permanent virtual circuit that has limited availability. That Type 1 interim systems also has a commercial equivalent. The previously mentioned government-sponsored IP encryptor will in fact produce a combined IP and ATM encryptor. Further government developments are being considered for tactical platforms and for end-system use.

In the commercial arena, two companies have produced ATM encryptors. One unit operates over DS-3 circuits to secure a single PVC. The other unit operates at 155 Mb/s and third unit operates at 622 Mb/s system due out in summer 1997. While none of these commercial units, Type 1 grade, this is an area for commercial investment consideration.

The incorporation of native mode firewalls in ATM is in early stages of demonstration. No Type 1 products incorporate that functionality at this time. Some commercial systems have been demonstrated that incorporate simple IP packet filters. It is expected that there would be a similar need for encrypting firewall technology in ATM networks just as there is in IP networks. Although some doubt the extensibility of good firewalls to the level of performance that would be required in an encrypting firewall application, practical network administration makes the near-term utility of such a device very attractive.

## Transport Layer Security

Over the last few years, more attention has been given to providing a set of common security services in end systems. One version that gained acceptance actually existed just above the transport layer and was called Secure Session Layer Security. This effort has migrated to the Internet Engineering Task Force and placed at the top of the transport layer. This service is being called Transport Layer Security (TLS). One advantage of TLS is that this is the first place in the network stack where security services can be broken out per application rather than applying generic services to a secure pipe. However, this set of security services must be implemented in end systems and is therefore subject to all the invocation concerns of application layer services. The traffic flow problem is even more acute in TLS because of the visibility of individual services. At this point only early commercial implementations of TLS exist and none of these are the equivalent of Type 1-grade standards.

## Super-encryption in VPNs

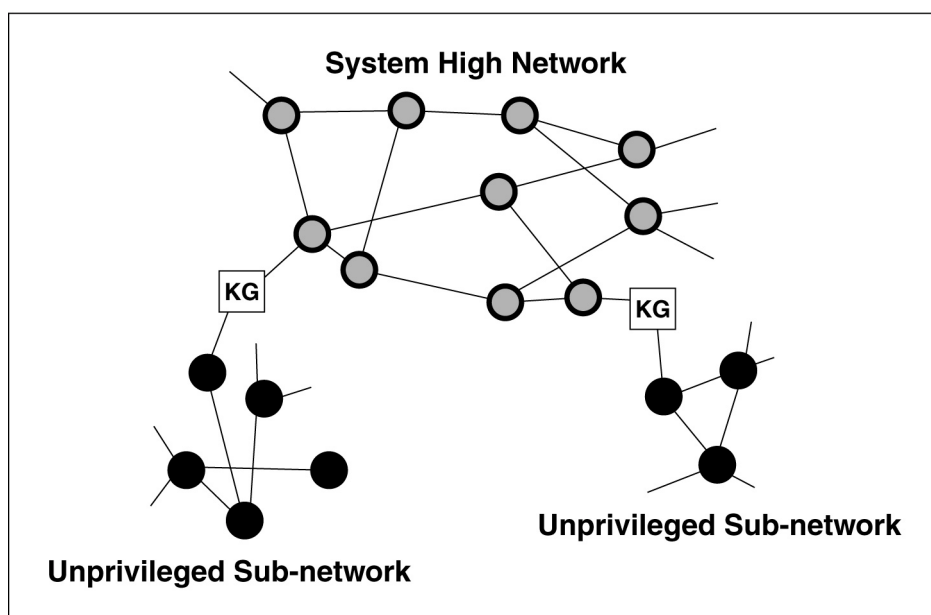
Super-encryption should be considered when there is a requirement to enforce privacy within the VPN. Such privacy may be implemented in end systems using lower assurance implementations of IPSec or ATM encryption under the control of a end system, TLS, or application-level mechanisms implemented either in hardware or software. Alternatively, an entire sub-network may be provided privacy using of a network encryption element. Note that this generalized description gives much flexibility to scale the level of protection mechanisms employed to fit the threat against an information system. Applicable architectures include link

protected switching and routing centers with end-system-based privacy mechanisms, link protected switching and routing centers with enclave-based privacy mechanisms, and enclave-based protection backed by end-system-based privacy mechanisms. For instance, one should consider link-protected switching and routing centers with network layer security mechanisms if there is a traffic flow security requirement and the switching centers are maintained by un-cleared personnel.

## Reverse Tunneling

In some scenarios, one needs to tunnel lower classification information through a higher classification system-high network. This is often accomplished by using the same high-grade cryptographic mechanism as would be required to tunnel high-grade traffic through a public network. Figure 5.3-4 illustrates the placement of cryptographic mechanisms for reverse tunneling.

The primary threat in this case is leakage of classified information into the lower classification tunnel. To help solve this problem, the cryptographic equipment should be under the control of the higher classification network and not under the control of the end users. If the lower level system is itself classified, it may have its own security mechanisms. It is recommended that the network layer confidentiality system use a tunnel mode rather than a transport mode mechanism if one is available. Tunneling maximizes the isolation between the levels of information and prevents the low side from using short cipher to elicit recognizable responses from nodes on the high side of the tunnel. Although it is traditional to use cryptography strong enough for protection of classified information in the reverse tunnel, the information within the tunnel may only be unclassified. An area for investigation is whether well-implemented commercial systems can be used for such applications. Good implementation must address the need for strong integrity mechanisms on the secure tunnel. This will help prevent malicious code within the VPN from infiltrating information through the lower level tunnel. Finally, the implementation should consider what, in analog radio frequency devices, would be called reverse isolation. In particular, careful attention must be paid to unintentional leakage of higher-level plaintext information through the encryptor and out the lower level information port.



**Figure 5.3-4 Reverse Tunneling Placement of Cryptographic Mechanisms**

Although it is traditional to use cryptography strong enough for protection of classified information in the reverse tunnel, the information within the tunnel may only be unclassified. An area for investigation is whether well-implemented commercial systems can be used for such applications. Good implementation must address the need for strong integrity mechanisms on the secure tunnel. This will help prevent malicious code within the VPN from infiltrating information through the lower level tunnel. Finally, the implementation should consider what, in analog radio frequency devices, would be called reverse isolation. In particular, careful attention must be paid to unintentional leakage of higher-level plaintext information through the encryptor and out the lower level information port.

## Relationship of Virtual Private Networking and Remote Access

The notion of virtual private networking implies an enclave of users who are protected from the network as a whole by some boundary device. Remote access implies a sole user gaining access to the enclave by some protected means. Although the mechanisms to implement this access may be similar to that used for VPN, the details of the connection are vastly different.

Although dial-up access through a phone line resembles a VPN implemented at layer 2, it can implement security mechanisms at layer 2 or layer 3. The preferable solution would be a layer 2 protection mechanism with strong user authentication. An acceptable solution would be a layer 3 IPSec solution, given that the AH is implemented in the solution and strong user authentication is required. What makes these solutions more acceptable is that data exchange occurs directly between end systems without the need for protocol negotiation with an untrusted entity.

Remote access through an Internet Service Provider (ISP) using IPSec resembles an IP-based VPN. The primary difference is that remote access through an ISP consists of a simultaneous connections to a private entity and a public entity without any intervening firewall or other protection mechanism. No monitoring of the information flow occurs between the remote host and the ISP to determine that no malicious transfers are taking place. This uncontrolled simultaneous connection between private and public entities takes this configuration outside the virtual private networking arena. Two areas of concern would have to be addressed before an ISP could be considered as a viable means of remote access to a secure enclave. The first concern is the window of unprotected access to the remote station during the period when the connection is made to the ISP but before IPSec or other mechanism can be invoked on communications with the secure enclave. The second is the concern that the remote terminal can become a convenient method for an insider to pass information outside the secure enclave because the remote terminal has simultaneous connection to the secure enclave and the unsecured ISP. The only solution would be a guaranteed invocation of the IPSec security mechanism across all IP source-destination pairs once a connection is made.

## Role of Firewall Technologies in VPNs

The resurgence of VPNs based on encryption mechanisms is largely due to concern about penetrability of firewalls. However, encryption alone will only create secure data pipes between enclaves. There are no restrictions on the type and content of information that can be carried by that pipe. Joining enclaves with a secure data pipe also creates a default security policy that is the sum of the most promiscuous aspects of the individual policies. There are many situations where this default policy applies. When are connecting peer entities where the primary threat to the information is from external sources and where either all personnel accessing the system possess the same level of clearance or they may be deemed so trustworthy that they would not access restricted information given the opportunity, secure data pipes alone may be sufficient security. If these assumptions are not valid, the secure pipes must be supplemented by additional separation mechanisms. Firewalls are one way of providing that additional separation. Appropriate firewalls can allow an administrator to control the types of

information flow across the VPN. For further discussion of firewall capabilities, see Section 6.1 (Firewalls).

It is important to reiterate that, in this case, the use of a firewall is recommended for the situation where two subnetworks are at the same security level but accreditors have assumed differing levels of risk in providing network security. Those interested in the case where high-to-low connections are required should refer to Section 6.3.1 (Guards) of this document.

There is a great diversity in the quality of implementation of firewall technology, and the purpose of this section is not to rate implementation quality. However, some general guidance on when to use firewalls and how restrictive they should be is appropriate.

- Primary protection between classified systems should be through some lower layer encryption system. Although these devices provide no protection against malicious users inside the network, it does limit accessibility of the VPN by outsiders.
- When true peers are connected, no firewall should be required.
- When applications demand high bandwidth, firewalls are likely to fail to meet the requirements. One area for suggested research is techniques to increase the throughput of a firewall while maintaining its effectiveness.
- When two connected systems are not exact peers, use of at least one firewall is recommended, and it should be placed at the enclave with the most demanding security requirements.
- When a firewall is required, the restrictions on connectivity should be commensurate with the minimum communications requirements and the difference between security levels and compartmentation within the respective enclaves.

## Interoperability of VPN Protection Technologies

Up to this point this section on VPNs is written as though the population were segmented into defined communities that have no communication with each other. Under these conditions, it is easy to define a unique security solution for each community. Within the DoD, such islands of communication cannot exist. During times of contingency, lines of communication are likely to be opened where none had been planned. This creates a conflict between the need for interoperability between organizations and the need to design a secure communications infrastructure that meets mission needs. The following are possible solutions to the interoperability problem.

- 1) Require a uniform communications and security infrastructure.
- 2) Require end systems to implement all security features and require peer-to-peer negotiations.
- 3) Implement gateways that convert information to plaintext and re-encrypt in the appropriate format.
- 4) Develop methods of maintaining confidentiality through interworking functions.
- 5) Implement redundant security mechanisms and modify protocol stacks to give visibility to the invocation of security mechanisms at all layers.



Of these options, 1 and 2 are unworkable for the following reasons.

- A uniform solution will not meet all requirements, and requiring that all systems carry all security mechanisms is too expensive.
- These options will likely result in failure to communicate if any of the peers fail to complete a secure setup, or in compromise if the default is to pass the requirement for securing the communications to the next higher layer when peers fail to negotiate secure setup.

Options 4 and 5 are research areas at this time. The TACLANE equipment, in some sense, is an early implementation of option 5. If a secure ATM call setup fails, the device assumes that communications must be secured via IPSec. This, however, is a point solution and does not address the breadth of interoperability problems.

Therefore, in the near term, the only viable solution is option 3, red gateways between dissimilarly protected networks. Research is needed to determine whether options 4 or 5 can be viable at some point in the future to reduce plaintext exposure created by the use of the option 3 red gateways.

## 5.3.6 Cases

To apply these security technologies to user networks, it is most convenient to describe typical situations that must be encountered. In each of these situations, it is assumed that the end networks are of a single level of classification, employ the same structure of components, and that consistent security policies are in place. The following cases are considered.

- 1) Classified networks connected over public infrastructures where indications and warnings are not a consideration.
- 2) Classified networks connected over public infrastructures where indications and warnings to adversaries must be considered.
- 3) Sensitive But Unclassified networks connected over public infrastructures.
- 4) Tunneling of lower classification information over a classified system-high network.
- 5) Tunneling of higher classification information over a classified network.
- 6) Maintaining compartmentation and privacy over a secured classified network.
- 7) Single backbone architectures for voice, video, and data.
- 8) Connection of networks where sub-networks have incompatible security policies.

## 5.3.7 Framework Guidance

### Case 1: Classified Networks Connected over Public Infrastructures Where Indications and Warnings Are NOT a Consideration

This case covers the connection of classified enclaves when traffic flow security is not a priority and it represents the majority of deployed classified VPNs. This case applies when the communications on the network are not involved in the planning and deployment of strategic or tactical forces, when the network is not involved in sensitive time-dependent operations, and/or when there is no tie to strategic intelligence sensors where reactions of the network can be used to probe the capabilities of sensors.

Three viable alternatives exist for creating secure VPNs over public infrastructures for this case. The most secure is to use Type 1 link layer protection; gives the greatest protection against outsider attacks on the network and the fewest means for malicious insiders to send information outside the network. This level of protection comes at the cost of increased communications cost and inflexibility in expanding or changing the network layout. Almost as good a choice would be to use circuit emulation on an ATM permanent virtual circuit using Type 1 ATM encryption, and this solution may give some cost flexibility.

If communication costs or the need for flexible communications precludes the use of leased circuits or circuit emulation, then network-layer-based solutions should be considered. Type 1 enclave-based solutions are recommended. NES is an example of a Type 1 enclave-based solution for IP-based network topologies. Other, more standardized Type 1 IPSec-compliant solutions also are available. FASTLANE and TACLANE provide Type 1 solutions for ATM-based topologies.

There are no host-based Type 1 systems for network layer protection at this time. While this class of solutions can potentially be very cost effective, the strength of invocation has not been sufficiently addressed to make a recommendation that such solutions be used. There are no commercial security systems of sufficient strength for protection of classified information at this time.

### Case 2: Classified Networks Connected Over Public Infrastructures Where Indications and Warnings to Adversaries MUST Be Considered

What distinguishes this case from the previous one is that observation of external traffic patterns even without decryption of the underlying information could give critical information to adversaries. For example, if a network extends into a tactical theater of operations, changes in traffic patterns may indicate the imminence of offensive operations thereby losing the element of surprise. Another example would be where a network can be identified as processing information from critical sensor platforms. Here probing the sensor and observing resulting traffic patterns can give away sensor response times and sensor sensitivity.

The basic solution set is the same as the previous case. The best solution is still the Type 1 link based security system. The reasons provided in case 1 still hold, with the addition of the complete traffic flow protection. Although, the existence of links can be easily identified, the change in traffic patterns is indiscernible.

If link-based solutions are not feasible, then prime consideration should go to enclave-based network layer solutions that tunnel multiple logical connections through a single path. This solution is represented by the NES because it tunnels enclave information via IP packets that are addressed from NES to NES. As Type 1 IPSec-compliant systems that use the IPSec tunnel mode become available, these systems also will meet security requirements. ATM wide area connections can provide some of the same capabilities for IP LANs because multiple IP source destination pairs can tunnel through the same ATM virtual circuit.

As end-to-end ATM applications become viable, tunneling will become more difficult because individual virtual circuits will be set up between end systems for each source-destination pair. The best solution for this case will be a secure virtual path service between enclaves which will at least enable identification of the end points of each virtual circuit to be encrypted within the virtual path. However, the characteristics of each data flow will be observable. When traffic analysis is a threat, any of the network-based solutions, especially the end-to-end ATM solution, can be made better with rate shaping of the traffic by the end systems.

No host-based Type 1 systems for network layer protection exist at this time. Although, this class of solutions can potentially be very cost effective, the strength of invocation has not been sufficiently addressed to allow a recommendation that such solutions be used. No commercial security systems of sufficient strength for protection of classified information exist at this time.

## Case 3: Sensitive But Unclassified Networks Connected Over Public Infrastructures

This case is the Sensitive But Unclassified (SBU) version of case 1. The difference in the solution is that commercial-strength mechanisms may be adequate for protection without going to the expense of a Type 1 equipment. The security benefit is probably insufficient to consider a link layer protected solution for the SBU case. It is recommended that a commercial enclave based network layer solution be used whether that solution is ATM or IP based. A mode that supports IPSec tunneling or the ATM equivalent, is preferable to a transport mode solution.

Host-based solutions are not recommended for primary protection of a direct connection to public networks until further testing has been accomplished to check strength of invocation and their ability to be bypassed.

## Case 4: Tunneling of Lower Classification Information Over a Classified System-High Network

This case exists when a classified network that already exists is protected at a link layer is used to transport unclassified or SBU information as a matter of convenience. In this situation protection has traditionally been implemented with Type 1-encryption systems as in the case of the tunneling of unclassified information through Secret Internet Protocol Router Network (SIPRNET) using the NES.

The properties desired from such a solution are that the mechanism be sufficient to protect of information that is presented to the network, that invocation cannot be bypassed, and that reverse isolation of the mechanism be sufficient to prevent leakage of the higher classification information onto the lower classified network. Strong data integrity mechanisms must be part of the security services offered by the security device used. These mechanism are used to protect the information on the low side of the connection but to eliminate the possibility of malicious insiders using the channel as a means to send information out of the secure network.

Although Type 1 solutions can still be used for such applications, commercial network layer systems should be considered. In addition, a tunneling mechanism should be mandatory. Note that this requirement eliminates IPSec transport mode solutions. The equipment implementing the security should be under the ownership, control, and configuration of the higher classification network. The system must not be able to be configured from the port that is connected to the lower classification network.

## Case 5: Tunneling of Higher Classification Information Over a Classified Network

An example of this type of application would be the tunneling of a Top-Secret network like the Joint Worldwide Intelligence Communications System (JWICS) through the Secret network SIPRNET.

The central issue in this case is whether the solution must be as strong as that required for tunneling over an unclassified network or, because protection is provided in case 4 to deal with the use of a lower classification network, whether a weaker mechanism can be considered.

It is recommended is that a Type 1 enclave-based tunneling mechanism be required. The mechanism should be under the control of the higher classification network.

## Case 6: Maintaining Compartmentation and Privacy Over a Secured Classified Network

The difference between this case and case 5 is that compartmentation is an enforcement of need-to-know among people who are equally cleared. It is assumed that the protection on the network is already sufficient to deter penetration by outsiders. Therefore the real need is for privacy within the network rather than protection from malicious outsiders. Although application layer solutions are sufficient for lower bandwidth applications, more demanding applications will probably require some network-based privacy solution.

Given the threat environment, this is an ideal case for using commercial host-based solutions, whether IP transport mode or ATM end-to-end.

## Case 7: Single Backbone Architectures for Voice, Video, and Data

This architecture was one of the primary motivations for the development of secure ATM (in addition to the scalability and the speed of implementation. By placing the security at the ATM

layer, a single set of mechanisms successfully protect all information that crosses an enclave boundary. That vision is too optimistic. Problems occur with voice connectivity. A secure voice architecture currently covers all transport means except broadband voice. Although ATM security is perfectly capable of protecting voice communications, the problem is the lack of secure interworking between broadband voice and secure N-ISDN and POTS voice. Until these interworking issues are resolved, it is not recommended that broadband voice services be secured with native mode ATM security services.

## Case 8: Connection of Networks Where Sub-networks Have Incompatible Security Policies

The previously recommended solutions for VPNs all assume that the enclaves have compatible security policies. Under present security guidelines and as a risk management philosophy becomes more widespread, security policies are likely to diverge. Therefore it is expected that enclaves to be connected will have security policies are incompatible in some way. In the standard virtual private networking scenario, the unimpeded flow of information within the virtual network create a resultant security policy that is a fusion of the most liberal aspects of the security policies of the individual enclaves. The system security administrators of the individual enclaves either need to recognize the resultant security policy and assess the impact on their systems or an additional separation mechanism must be added to help enforce the desired policy. This case is an ideal place for the marriage of firewalls with VPNs. In this respect, the commercial community is far ahead of the Type 1 community with the widespread availability of encrypting IPSec-compliant firewalls. When additional separation is required, an appropriate IP or ATM-based firewall that implements features needed by the enclave, cascaded with the Type 1 enclave protection mechanism is recommended.

## References

---

1. Bellovin, Steven M., Problem Areas for the IP Security Protocols, July 22-25, 1996, San Jose, CA: Proceedings of the Sixth Usenix UNIX Security Symposium, 1996  
<<http://www.usenix.org/publications/library/proceedings/sec96/bellovin.html>>

This site provides an abstract of the document. You must become a member of USENIX to see the full text of the document. To become a USENIX Member, see the Membership Information link on the Web site.

## Additional References

---

- a) Virtual Private Networks, Faulkner Information Service, Pennsauken, NJ, May 1996

## 5.4 Secure Voice

This section will be provided in a later release of the Framework.

This page intentionally left blank.



## 5.5 Multiple Security Layers

Users are struggling to implement networks in which information of different classification levels are being transported over the same backbone. Users are using need to know to create communities of interest. The network is being relied upon to provide data separation for each compartment. Guards that allow information to migrate from one compartment to another is a technology gap. Labels at the network layer, Closed User Groups (CUGs), and encryption are all technologies being investigated to provide reliable data separation. A new section to be supplied in a later release of the Framework.

This section will be provided in a later release of the Framework.

This page intentionally left blank.

## Chapter 6

# Defend the Enclave Boundary/ External Connections

An enclave is an environment under the control of a single authority with personnel and physical security measures. Enclaves typically contain multiple Local Area Networks (LANs) with computing resource components such as user platforms; network, application, and communication servers; printers; and local switching/routing equipment. This collection of local computing devices, governed by a single security policy, regardless of physical location is considered an “enclave.” Because security policies are unique to the type, or level, of information being processed, a single physical facility may have more than one enclave present. Local and remote elements that access resources within an enclave must satisfy the policy of that enclave. A single enclave may span a number of geographically separate locations with connectivity via commercially purchased point-to-point communications (e.g., T-1, T-3, Integrated Services Digital Network (ISDN)) or using Wide Area Network (WAN) connectivity such as the Internet.

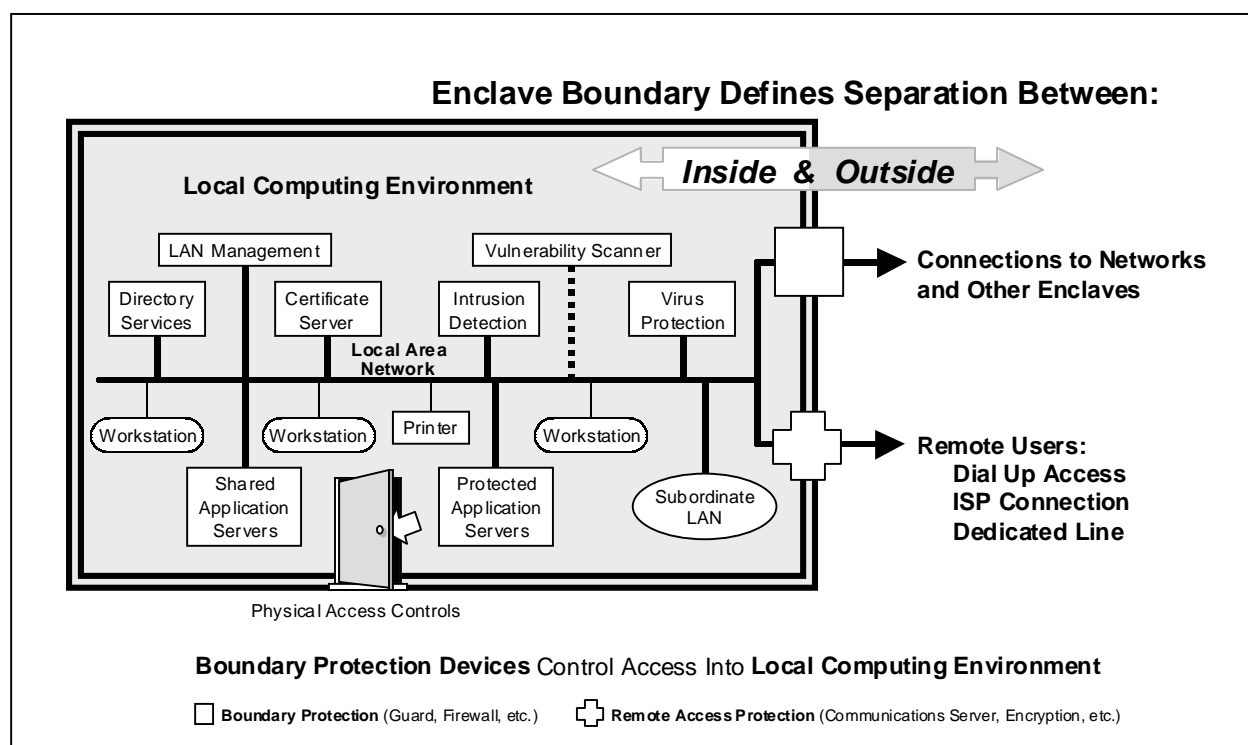
The majority of enclaves have external connections to other networks. These external connections may be single level connections, where the enclave and connected network are at the same privacy level. Or the connection may be a high-to-low/low-to-high transfer, where the enclave is at a higher or lower level than the connected network. Enclaves may also have remote access connections to traveling users or users located in remote locations. The point at which the enclave’s network service layer connects to another network’s service layer is the enclave boundary. Figure 6-1 highlights the enclave boundary target environments within the high-level information infrastructure context. The placement of boundary protection mechanisms in Figure 6-1 is notional; only representing suggested, not necessarily actual, placement of Information Assurance (IA) components.

Defense of the enclave boundary is focused on effective control and monitoring of data flow into and out of the enclave. Effective control measures include firewalls, guards, Virtual Private Networks (VPNs), and Identification and Authentication (I&A)/access control for remote users. Effective monitoring mechanisms include network-based Intrusion Detection System (IDS), vulnerability scanners, and virus detectors located on the LAN. These mechanisms work alone as well as in concert with each other to provide defenses for those systems within the enclave that cannot defend themselves or could be undermined by failures in systems operating at lower security levels or with less stringent security policies. Although the primary focus of the perimeter is on protecting the inside from the outside, enclave boundaries also provide some protection against malicious insiders who use the enclave to launch attacks or who facilitate outsider access through open doors/covert channels.

The IA strategy for defending an enclave boundary includes a number of general defensive measures, and specific capabilities that address remote access and interoperability across security levels. In general, the enclave perimeters must be established and be equipped with

professionally managed electronic access portals that enable effective control and monitoring. These portals should enable dynamic throttling of services in response to changing Information Conditions (INFOCONs). They should establish mandatory Department of Defense (DoD) policy on protocols that are allowed and disallowed between secure enclaves and external systems.

The strategy mandates the use of basic intrusion detection for all DoD enclaves, with additional detection mechanisms for mission critical and mission essential enclaves. VPNs, used to establish Community of Interests (COIs) (or Intranets) will not be used between enclaves that provide different degrees of security, unless other adequate measures are used to protect the stronger enclave from the weaker one. An important strategy consideration is that Detect capabilities not be lost with the increased use of encryption. This requires that protection and detection capabilities be planned together. For VPNs, the DoD strategy is to install the VPNs in such a way that network-based monitors can be placed on their clear-text side.



**Figure 6-1 Defend the Enclave Boundary**

Within the IA strategy, systems and enclaves that are provided remote access to a secure enclave must comply with the same security policy as the secure enclave, comply with approved remote access protocols, be authenticated at the enclave perimeter, and ensure that the entire secure enclave is not jeopardized by overrun of remote access points. In all cases, remote access will require authentication using approved techniques. At a minimum, this will use non-reusable passwords, preferably in encrypted form, or public key-based approaches.

Continuous authentication (vice authentication only at the beginning of a session) is preferred. For interoperability across security levels, the DoD infrastructures will be based on a multiple security level strategy in which separate system and network infrastructures are maintained at

each security level. The use of devices that control data transfers across security levels will be minimized. When required by operational necessity, these shall be implemented by an official Secret and Below Initiative (SABI) (or Top Secret and Below Initiative (TSABI)) process. High side servers that serve as gateways to receive low-to-high transfers will use operating systems that are capable of enforcing user-level access controls, are properly configured and operated using the concept of least privilege and include other appropriate layers of protection (to include tripwires for malicious software protection, pre-placed forensics, reporting of incidents and anomalous activity and host-based auditing).

This section of the framework addresses the role of IA technologies in providing protection for the enclave boundary for the following.

- Firewalls.
- Guards.
- Remote Access.
- Virus/malicious Code Detection.
- Intrusion Detection.
- Multi-Level Security.

**UNCLASSIFIED**

Defend the Enclave Boundary/External Connections  
IATF Release 3.0—September 2000

**This page intentionally left blank**

**UNCLASSIFIED**

## 6.1 Firewalls

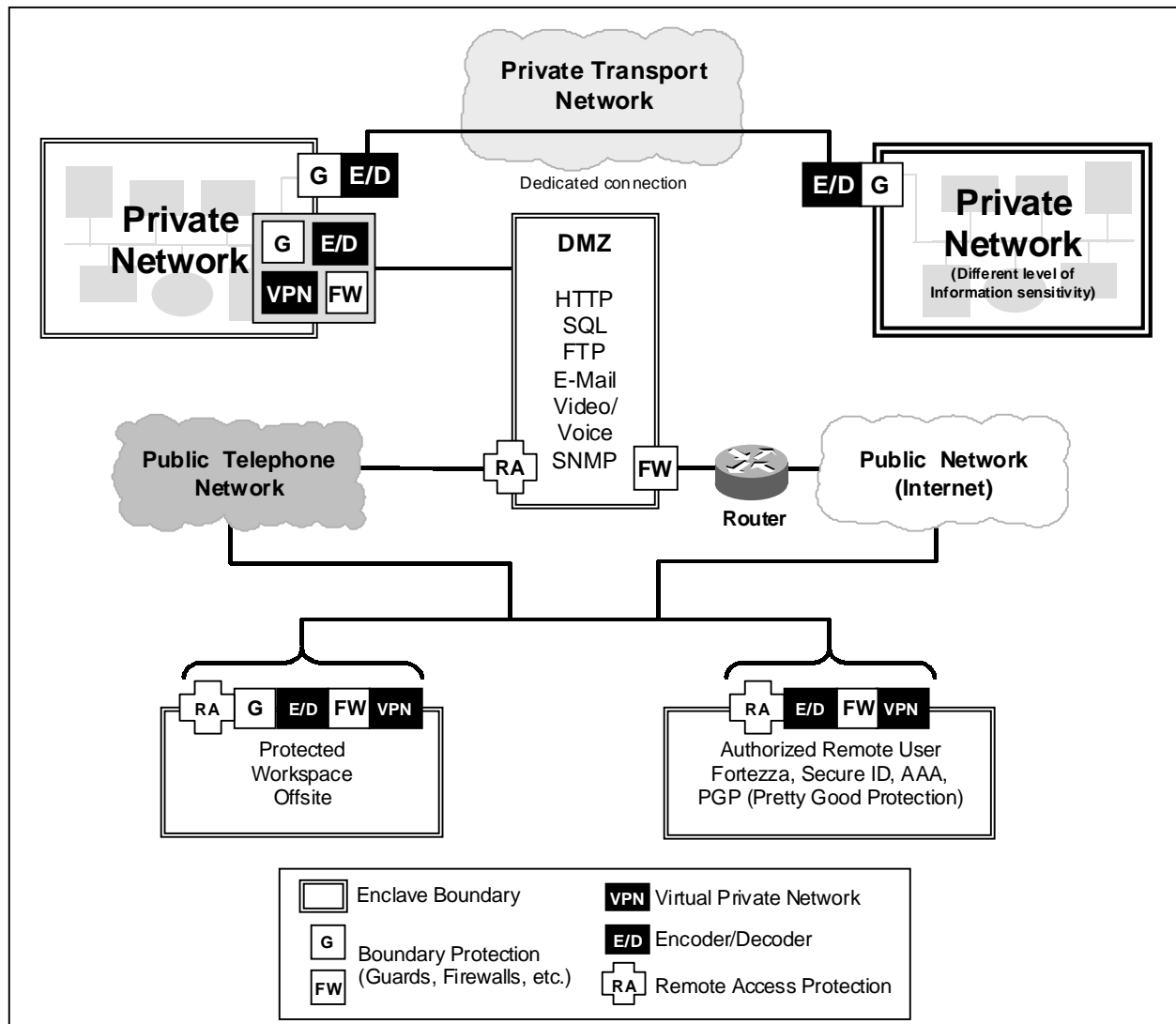
The purpose of a firewall is to protect internal information systems from external attacks. Firewalls address the requirement for authorized Local Area Network (LAN) users and administrators as well as individual workstation or personal-computer users, to safely access and be accessed-by untrusted (potentially hostile) external network connections. This means that all components inside the enclave boundary are protected against intrusion attacks: unauthorized extraction, modification, or deletion of data; denial-of-service; and, theft of resources or services. This firewall section addresses all components used for protecting interconnected, digital-electronic processing, transmission, or storage of information.

The focus of this Firewall section is on **external** electronic intrusions through the enclave boundary into a LAN or workstation that may be possible due to electronic connections. Attacks such as those performed by insiders or passive intercepts of traffic traversing backbone networks are not directly addressed within this section of the Information Assurance Technical Framework (IATF). While the unique concerns of the other protection categories are primarily addressed elsewhere in the Framework, there are some fundamental protection countermeasures—common to most environments—addressed here. Clearly, the concerns and approaches relevant to external electronic intrusions are interdependent with those of other protection categories (such as remote access, system high interconnects, Multi-Level Security (MLS), or security for applications). Thus, the following firewall-focused sections are intended to be complementary and integrated rather than separate, distinct layers of protection. For further expansion on site security please refer to <http://www.ietf.org/rfc/rfc2196.txt?number=2196>, RFC 2196, Site Security Handbook.) [1]

### 6.1.1 Target Environment

Users within an enclave can access external information services via network connections, dedicated connections, or dial-up connections. The environment illustrated in Figure 6.1-1 includes various combinations of methods of access involving Internet Service Providers (ISPs), Integrated Services Digital Networks (ISDN), Public Switched Telephone Networks (PSTN), X.25 Packet Exchange, wideband (cable-modems) and Internet and intranet networks/hosts that consist of both valid (trustworthy) agents and potentially hostile agents.

Included are those involving multiple access levels such as a private corporate LAN connecting to a public Wide Area Network (WAN), or a private corporate LAN connecting to a corporate intranet. The boundary protection approaches should be applied to many of the cases described in other categories (e.g., remote access, system high interconnections and Virtual Private Networks (VPNs)). Whenever networks (workstations) are interconnected, the Network Security Policy should require protection at the network access points, i.e. the enclave boundaries. Generally, the amount of protection needed increases as the sensitivity of the information increases, as differences in sensitivity levels increase, as the threat increases, and as the operational environment changes (likelihood for attack increases for high profile organizations).



**Figure 6.1-1 Enclave Boundary Environment**

## 6.1.2 Firewall Requirements

### 6.1.2.1 Functional Requirements

The following have been identified as representative ideal requirements based on a customer's perspective of needs.

- The user, if authorized, should have maximum access to needed information and services available on the WANs using any of the existing and emerging networking technologies and applications.
- The user and users' system should be protected against the full range of network attacks, be able to locate the source and type of intrusions, be able to react to such



intrusions, and be able to fully reconstitute the system following damage caused by intrusions.

- The approaches used to protect network access points should have minimal operational impact on the user.
- The approaches used to protect network access points should have minimal operational impact on performance of the associated components and networks.
- The approaches used chosen to protect network access points should be a scalable solution to allow for future needs.

## 6.1.2.2 Boundary Protection Mechanism Requirements

Boundary protection mechanisms are used to limit access to the internal network and are provided through the use of some combination of routers, firewalls, and guards. Please refer to Section 6.1.4.1, Technical Countermeasures, Boundary Protection via Firewalls, for further expansion on this subject. The following are typical requirements that boundary protection mechanisms should offer.

- Restrict sources, destinations, and services and block dangerous protocols such as Internet Control Message Protocol (ICMP). Both incoming and outgoing communications should be restricted.
- Restrict executable services and download capabilities.
- Employ internal Access Control Lists (ACL) where appropriate.
- Use Identification and Authentication (I&A) mechanisms—to include the use of software or hardware tokens—to authenticate outsiders to the boundary point.
- Use encryption to prevent interception of data which could provide the attacker with access into the network and for access control. This should include the encryption of remote management data.
- Hide the internal network (addresses, topology) from potential attackers using a mechanism such as network address translation.
- Log and analyze source-routed and other packets and react to or restrict attacks.
- Scan for malicious software.
- Facilitate proper boundary protection configuration by operators, e.g., user-friendly Graphical User Interface (GUI).
- Be self-monitoring and be capable of generating alarms.

Note that the intent of several of these countermeasures is to eliminate vulnerabilities of services that may not be needed by a particular user system. Current technologies do not permit complete user access to all desired services and destinations while simultaneously blocking all attacks. In addition, the use of encryption and certain identification & authentication mechanisms (such as hardware tokens) limits interoperability. Tradeoffs must be made.

### 6.1.2.3 Interoperability Requirements

The boundary protection should not force users to employ any nonstandard protocols or modes of operation nor any procedures that would prohibit interoperability with those external users or systems that users desire to communicate with and are permitted by the organization's network security policy.

- The firewall command and control channel must be secure to prevent eavesdroppers from learning the rules, Media Access Control (MAC) secrets, and other controlling data communicated over the firewall command and control channel (i.e., Simple Network Management Protocol (SNMP), Remote Monitor (RMON), Application Program Interface (API), and TELNET).
- An authentication mechanism is needed to prevent unauthorized entities from changing the rules. In the simplest case, IP-addressed based authentication may be satisfactory. If end-devices are allowed to modify the rules (as they are with SOCKS), then secure user-based authentication would have to be deployed along with an administration policy. For example, the policy may permit authenticated user A to open pinholes from his host at "high" port numbers and deny anything else. (SOCKS is out of the scope of this chapter, for more information please refer to <http://www.socks.nec.com> and <ftp://ftp.nec.com/pub/socks/>). [2, 3]

### 6.1.2.4 Anticipated Future Requirements

The approach employed to protect network access should allow for the evolution and reconfiguration of the network and associated components. The chosen approach should be scalable to allow for the future evolutions.

## 6.1.3 Potential Attacks

As previously stated, the focus of this firewall section is external attacks into a LAN or workstation that may be implemented by virtue of its electronic connections through the enclave boundary. The types of attacks are discussed below: active based attacks, distribution attacks, and insider attacks. Other attack categories (passive attacks and close-in attacks) are not directly addressed within the remainder of this chapter, but relate to this category and the technologies discussed. Please refer to Section 4.2, Adversaries, Threats (Motivations/Capabilities), and Attacks, for additional details regarding these attacks. Please also refer to Section 5.3, System High Interconnections and VPNs, for information on virtual private networking capabilities regarding security and protecting enclave assets from attacks.

### 6.1.3.1 Active Attacks

Attacks at the network access points generally fall within the active attacks category as defined in Section 4.2.1.4, Categories of Attacks. This type of attack has also been referred to as a "active attack". Any attempt to gain unauthorized access into a network or break network security features is an active attack. For more description please refer to Section 4.2.1.4.2, Table 4-2, Examples of Specific Active Attacks. Listed below are various examples of active attacks.

- Trick the Victim (Social Engineering).

- Masquerade as Authorized User/Server.
- Exploit System-Application and Operating System Software.
- Exploit Host or Network Trust.
- Exploit Data Execution.
- Exploit Protocols or Infrastructure Bugs.
- Denial of Service.

### 6.1.3.2 Distribution Attacks

Distribution attacks are the hostile modification of hardware or software. Such attacks can occur anytime hardware or software is transferred. For additional information, please refer to Section 4.2.1.4.4, Hardware/Software Distribution Vulnerabilities and Attacks and Table 4-3, Examples of Specific Modification Attacks. The following are examples of distribution attacks.

- Attacks to a system can occur via software distribution computer disks that are transferred among firewalls.
- An attack can occur when software is downloaded from the Internet, e-mail, or an internal LAN system.
- Attacks can occur through modifications made to hardware or software at the factory before distribution or during distribution. Malicious changes to software code or malicious modification of hardware can occur between the time it is produced in the factory and the time it is installed and used.
- An attack can also occur while a firewall is being configured, especially if being configured from remote locations.

### 6.1.3.3 Insider Attacks

Although the emphasis of protecting network access points is on protecting the inside from a potentially hostile outside world, mechanisms are needed for protection against outside and inside intruders. Thus, some of the technologies identified in this section apply to both insider and outsider threats. Further, once an outsider has successfully attacked a system to obtain access, the outsider, in effect, maneuvers within the system as an insider would. Technologies such as those designed to detect attacks by an insider may in fact be used in a similar manner to detect outsider attacks.

Insider attacks can occur when an **authorized user** (i.e., a person who has authorization into the system) remotely connects to the system and unintentionally causes damage to the information or to the information processing system. This nonmalicious attack can occur either from the user not having the proper knowledge or by carelessness. Malicious insider attacks are those in which an **authorized user** causes damage to the system or enters areas where the user is not authorized. Malicious attacks can also be caused by an unauthorized individual employing an authorized user's Personal Computer (PC) to maneuver within the system and cause damage. An example would be when an authorized user's laptop computer is stolen and then used to gain access into the system. For more information, please refer to Section 4.2.1.4.3, Insider Vulnerabilities and Attacks.

## 6.1.4 Potential Countermeasures

Fundamentally, protecting network access points from potential attacks can be addressed by limiting access to and from the LAN or workstation. In the protection of a network; important issues that need to be addressed include detecting and identifying malicious or non-malicious insider attacks, identifying potential vulnerabilities and attacks that may occur given the current configuration; and responding to, deterring, and recovering from detected attacks. The following subsections describe security requirements applicable to addressing attacks through an enclave boundary. Several of the countermeasures are covered in detail within other IATF focus areas and are listed as applicable. The countermeasure requirements are grouped under the two primary headings of Technical Countermeasures and Administrative Countermeasures.

### 6.1.4.1 Technical Countermeasures

#### Boundary Protection via Firewalls

Connecting through the enclave boundary to external resources such as the Internet introduces a number of security risks to an organization's information and resources. The first step in minimizing those risks consists of developing a comprehensive network security policy. This network security policy framework should include firewalls as boundary protection mechanisms. Boundary protection mechanisms can provide a measure of protection for a network or an individual workstation within the enclave boundary. The boundary protection device is intended to operate primarily as an access control device, limiting the traffic that can pass through the enclave boundary into the network. In general, boundary protection is provided through the use of some combination of routers, firewalls, and guards. Please refer to Section 6.1.1.2, Firewall Requirements, Boundary Protection Mechanism Requirements for additional information.

Although the main focus of this section is firewalls, a definition of routers and guards follows. A router that is configured to act as a firewall is a packet-filtering device that operates at multiple layers and permits or denies traffic through the enclave boundary into the internal network based on a set of filters established by the administrator. A guard is generally a highly assured device that negotiates the transfer of data between enclaves operating at different security levels. Please refer to Section 6.3, Guards, for more information. In contrast, a firewall is a boundary protection device between networks that are communicating at the same security level.

A firewall is a collection of components placed between two networks (or an individual workstation and a network) with the following properties.

- All traffic from inside to outside and vice versa must pass through this mechanism.
- Only authorized traffic, as defined by the local network security policy, will be allowed to pass.
- The mechanism itself is immune to penetration.

The firewall is thus a tool for enforcing the network security policy at the enclave boundary and has several distinct advantages as a protected network access device. First, the firewall allows

for centralized network security management, as it becomes the focal point for network security decisions. In addition, as the only directly accessible component of the enclave network, the firewall limits the exposure of the network to attack. By implementing and following a well-defined network security policy, maintaining cognizance of current vulnerabilities, reviewing audit data, and using available scanning tools, the security of the enclave is greatly enhanced.

However, there are disadvantages to using firewalls. They can be the single points of attack to the enclave. Firewalls do not protect the network and workstations within the enclave against most data-driven attacks, some denial-of-service attacks, social engineering attacks, and the malicious insider. Firewalls can thus potentially provide a false sense of security. Firewalls must be looked at as being only one part of a larger network security approach.

## Access Constraint

Measures that should be taken to constrain access to facilitate defense of enclave boundaries include the following.

- Provide data separation. For data that is allowed access to the protected network or workstation, steps should be taken to constrain as much as possible the amount of the system that can be affected. Steps that could be taken include allowing executables to run only in a particular domain or only on a server reserved for such purposes as discussed in Section 6.3, Guards.
- Employ application-level access control. Access restrictions may also be implemented within the enclave—within workstations or at various points within a LAN—to provide additional layers and granularity of protection. See Access Control List under Section 6.3.5.3, Processing, Filtering, and Blocking Technologies.
- Provide authenticated access control and (as appropriate) encryption for network management. See a previous subheading in this category, Boundary Protection via Firewall and Section 6.3.5.1, Authenticated Parties Technologies.

### 6.1.4.2 Administrative Countermeasures

While defending the enclave boundary, administrative countermeasures should be implemented with the boundary protection mechanisms and throughout the enclave. Quality network management and network security administration are imperative in maximizing the security of the network's configuration and protection mechanisms and increasing the likelihood of detecting vulnerabilities and attacks. For additional detail on security system administration, see Section 4.4.7, Administrative and Section 4.4.8 Stuff People Do. The following administrative mechanisms act as countermeasures to the various attacks mentioned in the Section 6.1.3, Potential Attacks.

- Be prepared for severe denial-of-service attacks, e.g., institute and practice contingency plans for alternate services.
- Routinely inspect for physical penetrations of the firewall.
- Educate users and staff on correct procedures when dealing with firewalls.
- Institute and exercise well-publicized firewall procedures for problem reporting and handling.

- Institute and exercise suspicious behavior reporting channels.
- Institute and monitor critical access controls, e.g., restrict changeable passwords, require dial-back modems.
- Minimize use of the Internet for mission or time-critical connectivity.
- Require security-critical transactions—e.g., establishing identity when registering—to be conducted in-person.
- Use trusted software where available and practical.
- Use subversion-constraining software and techniques wherever possible, e.g., avoid software that uses pointers which could be employed by a software developer to access unauthorized memory locations.
- Carefully map relationships between hosts and networks, constraining transitive trust wherever possible.
- Minimize cross sharing between users and file systems, particularly for high-sensitivity or high-threat applications, allowing only essential functions that have compelling justifications for sharing.
- Where possible, do not rely on Domain Name Server (DNS) for security sensitive transactions where spoofing an Internet Protocol (IP) address could cause problems.
- Institute, exercise, and monitor a strict computer emergency response team alert and bulletin awareness and patch program.
- Institute and practice procedures for recovery from attack when the firewall is penetrated.

## Countermeasure Effectiveness

The following is a list of attacks and the most successful countermeasures against these attacks. More detailed information about the types of attacks is also provided in Section 4.2, Adversaries, Threats (Motivations/Capabilities), and Attacks.

**Trick the Victim (Social Engineering)**—The best defense against this type of attack is to educate system/network users. The users must be aware that attempts may be made to obtain their passwords to enable access into the network or into secure areas of the network that the attacker may not be authorized to access.

**Masquerade**—The best technical countermeasure against this type of defense is to identify and authenticate outsiders and to use access constraints to authenticate and encrypt data. Administrative countermeasures that have high levels of effectiveness include using and monitoring access controls and minimizing the use of the Internet for critical communications.

**Exploit Software Vulnerabilities**—The highest defenses against attacks made by exploiting vulnerabilities of software include subverting constrained software, monitoring the Computer Emergency Response Team (CERT), obtaining patches, and minimizing the use of the Internet for critical communications.

**Exploit Host or Network Trust**—Minimizing use of the Internet for critical communications and subverting constrained software provides the highest level of defense against attacks exploiting the host or trust in the network.

**Exploit via Executables**—Attacks against the enclave boundary through executable applications can be fought through technical and administrative countermeasures. Overall technical measures that can be implemented include boundary protection, access constraints, and detection mechanisms. Boundary protection offers the best technical defense by restricting sources and services, by restricting the ability to download, and by restricting executables. Administrative measures to counteract attacks via executables are minimizing the use of the Internet for critical communications and using subversion-constraining software.

**Exploit Protocol Bugs**—To protect against protocol bugs, the two countermeasures providing the best defense are—once again—minimizing the use of the Internet for critical communications and using subversion-constraining software.

**Denial of Service**—The best technical defense for a denial of service attack made against a system is to have a detection and response system in place. Administrative countermeasures include advance planning to be able to offer service alternatives, minimize Internet usage for critical communications, and to have documented and rehearsed recovery procedures in place to help reconstitute the system.

## 6.1.5 Firewall Technology Assessment

### Access Control/Filtering

Access control/filtering is the main function of every firewall. This function can be accomplished in several ways ranging from a proxy at the application layer of the Operating System Interconnection (OSI) model to stateful inspection at the IP layer. By its nature, the firewall implements a specific network security policy that corresponds to the level of sensitivity of the boundary the firewall is protecting. The main fundamental purpose of the security policy is to limit access to the network and systems inside the enclave boundary from external sources. Only necessary inbound connections and services should be allowed. The firewall also restricts the connectivity of internal users to external destinations. Although internal users are generally trusted, they should be limited in what services they can use through the firewall to prevent them from unintentionally opening security vulnerabilities. The different firewall technologies offer different granularities of access control. Some firewalls are now capable of what were traditionally guard-like filtering functions. For example, firewalls incorporate software that filters access to either specific Universal Resource Locators (URL) or categories of URLs. Certain File Transfer Protocol (FTP) commands can be blocked while other commands are allowed through the firewall. Technology will continue to develop in this area. Very sophisticated and highly refined access control capabilities are likely to become standard firewall features.

### Identification & Authentication

Identification and authentication is one of the major functions provided by the different firewall products. While users on the inside of a firewall, inside the enclave boundary, are often considered trusted, external users who require access to the internal network must be authenticated. Most security experts agree that passwords are not a strong method of authentication. In fact, cracking user passwords is one of the most common system attacks. Other authentication methods for screening access through a firewall include one-time passwords,

time-based passwords, and challenge-response schemes. The most common one-time password system in use is S\key, a software-based authentication mechanism using Message Digest 4 (MD4) or Message Digest 5 (MD5). S\key works by starting with a seed and applying MD4 or MD5 to generate a sequence of keys. S\key encodes the keys into a series of short words and prompts the user for the previous key, n-1, then S\key applies the MD4 or MD5 to the user's answer and checks to see if the result is the key n that it knows. Time-based passwords are a special form of one-time password. In these systems, the password varies at a specified time interval based on an internal algorithm, thus adding the additional complication of maintaining clock synchronization. Challenge-response systems are more complex and involve something the user has (a smart card or PC card) and something the user knows (password). Although it is possible to implement these systems in software, using hardware tokens has numerous advantages. Commercial firewall products support a wide range of authentication mechanisms.

## Mobile Code Blocking

In addition to more basic blocks of mobile code (Java, \* Script, ActiveX, etc.), firewall systems are beginning to offer containment for the execution of mobile code. This includes sandbox machines isolated from the rest of the network and restricted environments to run the Java Virtual Machine (VM) within. Reference RFC 1918 –Address Allocation for Private Internets for more information, <http://www.ietf.org/rfc/rfc1918.txt?number=1918> [4]

## Encryption

Firewalls become a focal point for the enforcement of security policy. Some firewalls take advantage of this to provide additional security services, including traffic encryption and decryption. To communicate in encryption mode, the sending and receiving firewalls must use compatible encrypting systems. Current standards efforts in encryption and key management have begun to allow different manufacturers' firewalls to communicate securely. To address this situation, vendors have been working on a network-level encryption interoperability approach through the Internet Protocol Security (IPSec) standard, set forth by the Internet Engineering Task Force (IETF). However, these efforts require further development before the customer can assume compatibility. Firewall-to-firewall encryption is thus used for secure communication over the Internet between known entities with prior arrangement, rather than for any-to-any connections. Verifying the authenticity of system users is another important part of network security. Firewalls can perform sophisticated authentication, using smart cards, tokens, and other methods.

## Auditing

Auditing refers to the tracking of activity by users and administrators. As opposed to accounting—where the purpose is to track consumption of resources—the purpose of auditing is to determine the nature of a user's network activity. Examples of auditing information include the identity of the user, the nature of the services used, what hosts were accessed when, what protocols were used, etc.



## Network Address Translation

Network Address Translation (NAT) is a method by which IP addresses are mapped from one realm to another to provide transparent routing to hosts. NAT enables a LAN to use one set of IP addresses for internal traffic and a second set of addresses for external traffic. Traditionally, NAT devices are used to connect an isolated address realm with private unregistered addresses to an external realm with globally unique registered addresses (Internet). That is, a NAT device sits at the enclave boundary between the LAN and the Internet and makes all necessary IP address translations.

## Resist Penetration

Another important aspect of a firewall is how well it protects itself against attack. The firewall itself should resist penetration, because breaking into the firewall will give a hacker access to the entire network. Most firewalls run on stripped-down versions of the operating system; unnecessary executables, compilers, and other dangerous files are removed. In addition, some firewalls employ technology that makes penetrating the firewall operating system extremely difficult. These firewalls are built on trusted operating systems or use mechanisms such as type enforcement (i.e., packet filtering, stateful packet filtering, proxy service, and application/circuit gateways) to provide this extra protection against penetration. Although these types of additional safeguards are traditionally found on guard devices, firewalls are also beginning to offer this type of extra protection against enclave boundary penetration.

## Configuration & Third Party Monitoring

Properly configuring the firewall components is critical to the security of the enclave boundary. Most vulnerabilities in firewalls arise from the improper configuration or maintenance of the firewall. For this reason, it is important to examine the administrative interface provided by the firewall. A GUI alone will not make the firewall any more secure. However, a well-designed operator interface can ease the administrative burden and more effectively illustrate how well the firewall has implemented the security policy. Firewalls also make use of various self-monitoring tools. These tools can provide additional access controls, can increase the auditing capability of the firewall, and can provide for an integrity check on the file system of the firewall. Some of these tools are proprietary and are provided with the firewall; other tools are available from the third parties and can be used to enhance the security of the firewall.

### 6.1.5.1 Firewall Types

#### Packet Filtering

Because routers are commonly deployed where networks with differing security requirements and policy meet, it makes sense to employ packet filtering on routers to allow only authorized network traffic, to the extent possible. The use of packet filtering in those routers can be a cost-effective mechanism to add firewall capability to an existing routing infrastructure.

As the name implies, “packet filters” select packets to filter (discard) during the routing process. These filtering decisions are usually based on comparing the contents of the individual packet headers (e.g., source address, destination address, protocol, and port) against preset rule

sets. Some packet filter implementations offer filtering capabilities based on other information beyond the header. These are discussed below in Stateful Pack Filtering. Packet filtering routers offer the highest performance firewall mechanism. However, they are harder to configure because they are configured at a lower level, requiring a detailed understanding of protocols.

## Stateful Packet Filtering

Stateful packet filtering technology, also referred to as *stateful inspection*, provides an enhanced level of network security compared to the static packet filtering described above. The stateful packet filter—working at layer 3 of the OSI model to examine the state of active network connections—looks at the same header information as packet filters do, but can also look into the data of the packet where the application protocol appears. Based on the information gathered, stateful packet filtering determines what packets to accept or reject. More importantly this technology allows the firewall to dynamically maintain state and context information about *past* packets. Thus, the stateful packet filter compares the first packet in a connection to the rule set. If the first packet is permitted through, the stateful packet filter adds the information to an internal database. This stored information allows subsequent packets in that connection to pass quickly through the firewall.

Network security decisions can then be based on this state information. For example, the firewall can respond to an FTP port command by dynamically allowing a connection back to a particular port. Because they have the capability of retaining state information, stateful packet filters permit User Datagram Protocol (UDP)-based services (not commonly supported by firewalls) to pass through the firewall. Thus stateful packet filters are advertised to offer greater flexibility and scalability. Stateful packet filtering technology also allows for logging and auditing and can provide strong authentication for certain services. Logging, or authentication as required by the rule set, occurs at the application layer (OSI layer 7). A typical stateful packet filtering firewall will log only the source and destination IP addresses and ports, similar to logging with a router.

Unlike application-level gateways, stateful inspection uses business rules defined by the administrator and therefore does not rely on predefined application information. Stateful inspection also takes less processing power than application-level analysis. However, stateful inspection firewalls do not recognize specific applications and thus are unable to apply different rules to different applications.

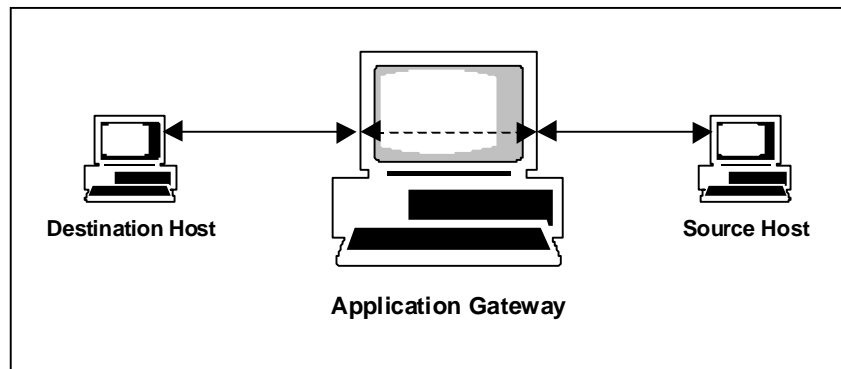
## Proxy Service, Application Gateways and Circuit Gateways

Figure 6.1-2, shows how proxy services prevent traffic from directly passing between networks. Rather, Proxy Services are software applications that allow for connections of only those application sessions (e.g. TELNET, FTP, DNS, Simple Mail Transfer Protocol (SMTP) for which there is a proxy. Thus, proxy services are *application-level* firewalls. The host running the proxy service is referred to as an application gateway. Since an application-level gateway is a system set-up specifically to counter attacks coming via the external network, it is also referred to as a bastion host. If the application gateway contains proxies for only TELNET or DNS, only these sessions will be allowed into the subnetwork. If a proxy does not exist on the application gateway for a particular session (TELNET, DNS, FTP, SMTP), those sessions will be completely blocked. Therefore, only essential services should be installed on the bastion host for if a serv-

ice is not installed, it cannot be attacked. Proxy services can also filter connections through the enclave boundary by denying the use of particular commands within the protocol session (e.g., the FTP put command) and by determining which internal hosts can be accessed by that service.

By using an application gateway through which access to the subnetwork is permitted, internal information can be hidden from systems outside the enclave boundary. The application gateway can provide a means for strong authentication by requiring additional authentication such as an additional password or

the use of a smart card. Each proxy contained within the bastion host can also be set up to require yet another password before permitting access. The bastion host as well as each proxy service is able to maintain detailed information by logging all traffic and the details of the connections. Logging helps in the discovery of and response to attacks. Each proxy is independent of all other proxies that may be running on the bastion host, so any operational malfunction of one proxy will not affect the operation of the other proxies. This also allows for ease of installation and removal of proxies from the system.



**Figure 6.1-2 Application Gateway**

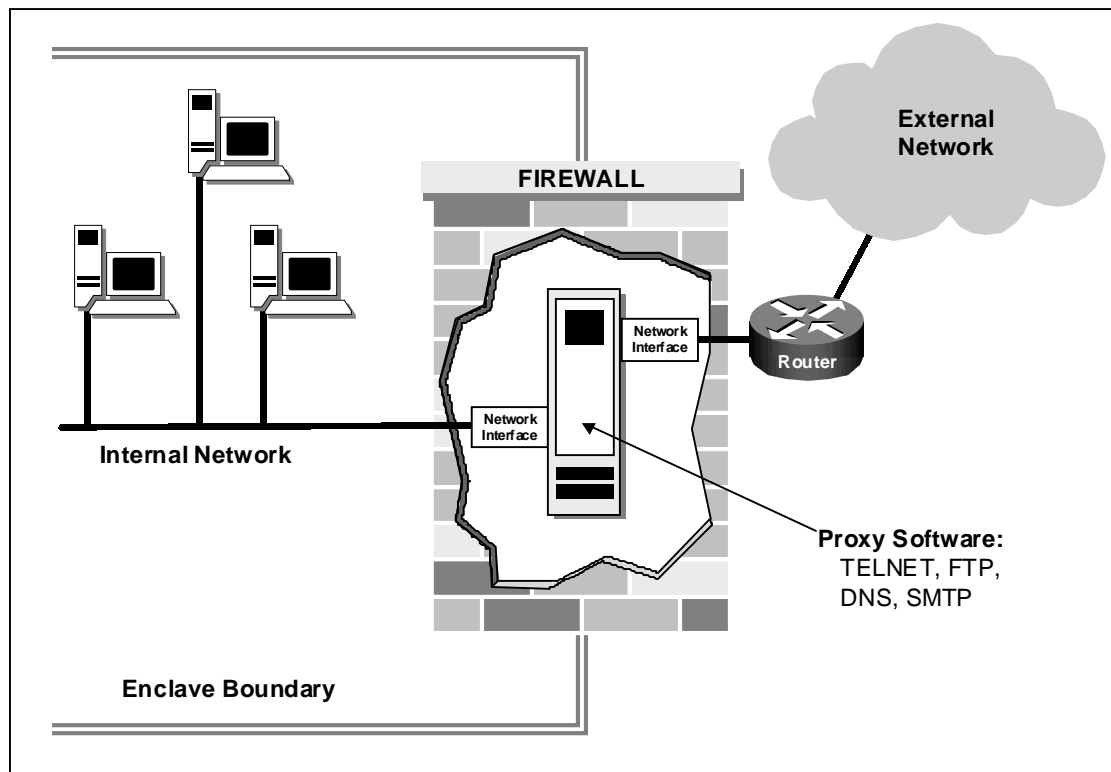
Circuit-level gateways are another type of firewall. A circuit-level gateway relays Transmission Control Protocol (TCP) connections without performing any additional packet processing or filtering. Circuit-level gateways are often used for outgoing connections where internal users are trusted. Outbound connections are passed through the enclave boundary based on policy, and inbound connections are blocked. Permission is granted by port address, upon which management control is primarily based. Although a circuit-level gateway is a function that can be performed by an application-level gateway, it is not as secure as an application-level gateway. When completing a connection, checking is not conducted to verify if application protocols (proxies) exist on the application gateway. Therefore, a circuit relay will not detect the violation if approved port numbers are used to run unapproved applications. A circuit-level proxy, acting as a wire, can be used across several application protocols. A bastion host can be configured as a hybrid gateway supporting application-level or proxy services for inbound connections and circuit-level functions for outbound connections. Circuit-level firewalls are less common than application-level firewalls due to the high probability that client modifications will be necessary to allow use of the circuit-level protocol.

Application gateways are generally dual-homed, which means that they are connected to both the protected network and the public network; however, they can be used in other configurations as discussed below. Packet filtering firewalls can also be dual-homed.

## 6.1.5.2 Firewall Architectures

### Dual-Homed

A type of an application-level firewall is the dual-homed gateway, which is a highly secured host that runs proxy software. A dual-homed gateway architecture has two network interfaces, one on each network, and blocks all traffic passing through it, as shown in Figure 6.1-3. That is, the host does not have the ability to directly forward traffic between the two interfaces. Bypassing the proxy services is not allowed. The physical topology forces all traffic destined for the private network through the bastion host and provides additional security when outside users are granted direct access to the information server.

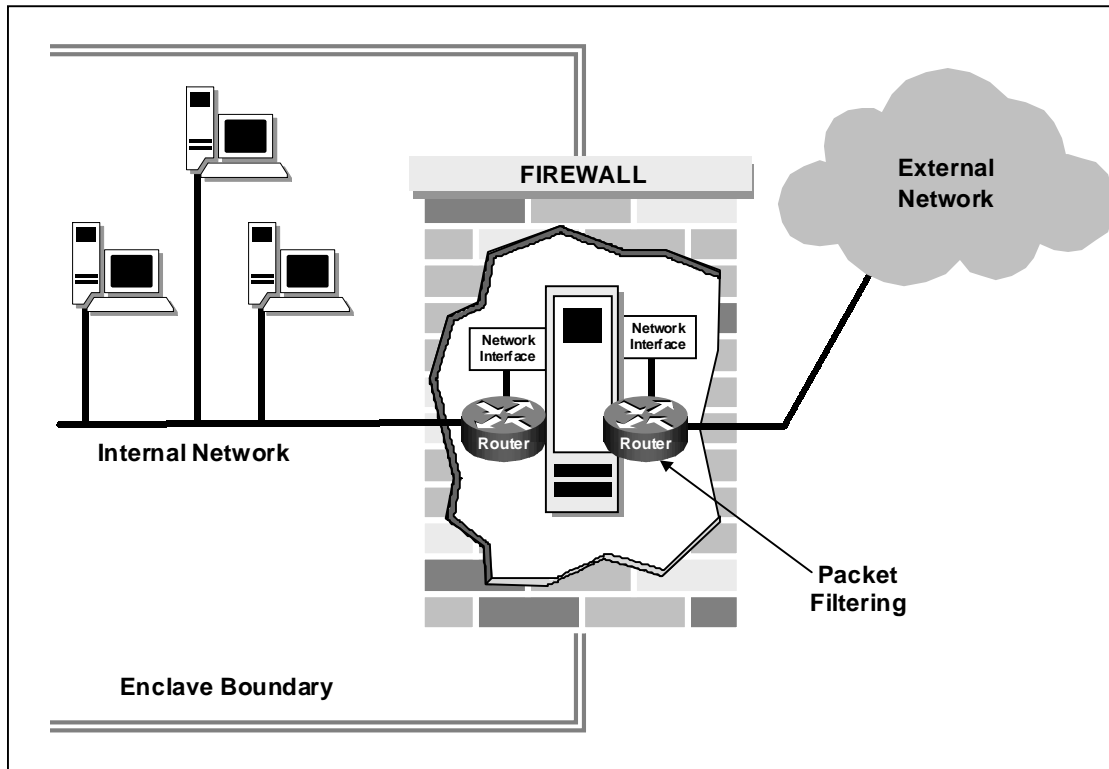


**Figure 6.1-3 Dual-Homed Firewall Architecture**

### Screened Host (Hybrid)

A screened host is a type of firewall that implements both network-layer and application-layer security by using both a packet-filtering router and a bastion host. A screened host architecture is also known as a hybrid architecture. This type of firewall architecture provides a higher level of network security, requiring an attacker to penetrate two separate systems. The system is set up with a packet filtering router sitting between an untrusted (external) network and the bastion host on the protected network so that only allowable traffic from untrusted networks pass to or from the internal bastion host, please see Figure 6.1-4. The packet filtering router is configured in such a manner that outside traffic has access only to the bastion host. An addi-

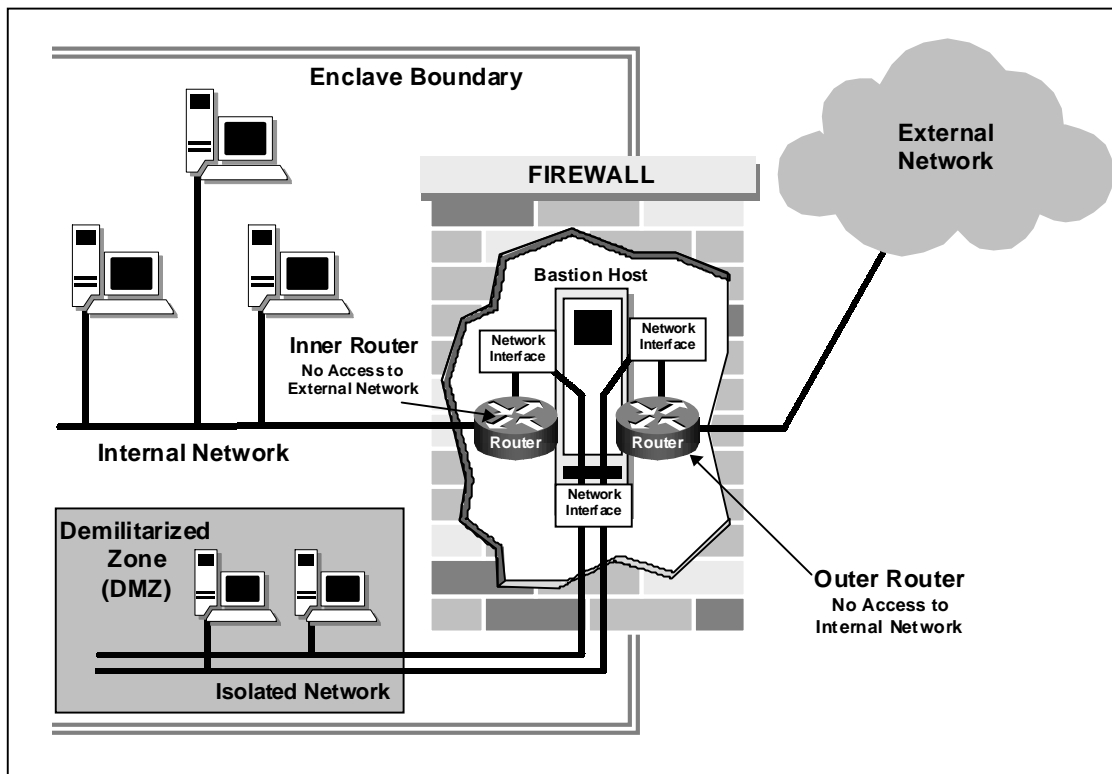
tional router may be set-up between the Bastion Host & the internal network for a greater level of security.



**Figure 6.1-4 Screened Host Firewall Architecture**

## Screened Subnet

Adding a small isolated network or Demilitarized Zone (DMZ) between the untrusted (external) network and the protected network creates a screened subnet, as shown in Figure 6.1-5. This small network is created with two routers. One router is considered the outer router and sits between this small isolated network/DMZ and the untrusted network. The function of the outer router is to restrict access from the untrusted network to specific systems within the screened subnet. This router also blocks traffic to the unprotected network originating from systems that should not be originating connections (e.g., information servers, site systems). The second router is considered the inner router and sits between the small isolated network/DMZ and the protected network. This architecture provides security redundancy, requiring an attacker to get past both of the routers before gaining access to the site systems. Even if an attacker broke into the bastion host, the inner router would still have to be penetrated before gaining access to the internal network. There is no single vulnerable point that will compromise the internal network.



**Figure 6.1-5 Screened Subnet Firewall Architecture**

The screened subnet firewall may be more appropriate for sites with large amounts of traffic or sites with high-speed traffic. A screened subnet can be made more flexible by permitting certain trusted services to pass from the external network to the protected network. But this may weaken the firewall by allowing exceptions. Greater throughput can be achieved when a router is used as the gateway to the protected subnet. Because routers can direct traffic to specific systems, the application gateway does not necessarily need to be dual-homed. However, a dual-homed gateway is less susceptible to weakening. With a dual-homed gateway, services cannot be passed for which there is not a proxy. The screened subnet firewall could also be used to provide a location to house systems that need direct access to services.

### 6.1.5.3 Firewall Selection Criteria

When selecting a firewall system the following should be taken into consideration.

- The firewall should be able to support a “deny all services except those specifically permitted” design policy, even if that is not the policy used.
- The firewall should support your network security policy, not impose one.
- The firewall should be flexible; it should be able to accommodate new services and needs if the network security policy of the organization changes.
- The firewall should contain advanced authentication measures or should contain the hooks for installing advanced authentication measures.

- The firewall should employ filtering techniques to permit or deny services to specified host systems as needed.
- The IP filtering language should be flexible, user-friendly to program, and should filter on as many attributes as possible, including source and destination IP address, protocol type, source and destination TCP/UDP port, and inbound and outbound interface.
- The firewall should use proxy services for services such as FTP and TELNET, so that advanced authentication measures can be employed and centralized at the firewall. If services such as Network News Transfer Protocol (NNTP), X Window System (X), HyperText Transfer Protocol (HTTP), or gopher are required, the firewall should contain the corresponding proxy services.
- The firewall should contain the ability to centralize SMTP access, to reduce direct SMTP connections between site and remote systems. This results in centralized handling of site e-mail.
- The firewall should accommodate public access to the site in such a way that public information servers can be protected by the firewall but can be segregated from site systems that do not require public access.
- The firewall should contain the ability to concentrate and filter dial-in access.
- The firewall should contain mechanisms for logging traffic and suspicious activity and should contain mechanisms for log reduction to ensure logs are readable and understandable.
- If the firewall requires an operating system such as UNIX, a secured version of the operating system should be part of the firewall, with other network security tools as necessary to ensure firewall host integrity. The operating system at start up should have all current and approved patches installed.
- The firewall should be designed and implemented in such a manner that its strength and correctness is verifiable. It should be simple in design so that it can be understood and maintained.
- The firewall and any corresponding operating system should be maintained with current and approved patches and other bug fixes in a timely manner.

## 6.1.6 Cases

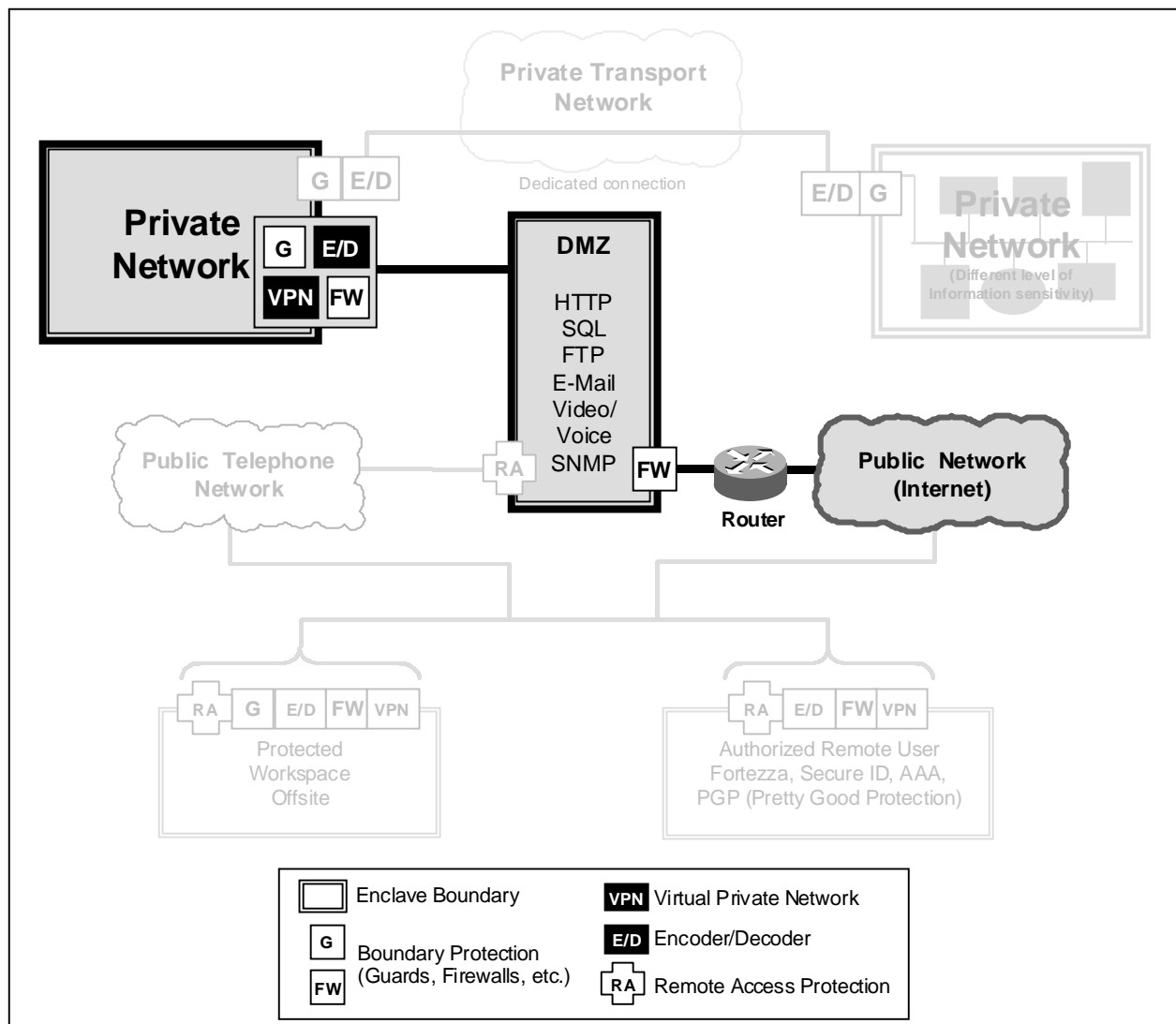
### Case 1

***A user communicating from a protected network to a public network. The information that is being sent is unclassified but private.***

This case is of the typical user connecting and passing information across the Internet. In Figure 6.1-6, a workstation within the protected network is communicating with the Internet. When connecting to a network of a lower protection level, mechanisms should be in place at the enclave boundary to provide protection for the users' workstation and the protected network.

A firewall can be deployed as part of an effective Boundary Protection function. Other components of Boundary Protection that can be implemented are through email, browsers, operating

system configuration; and router configuration. Once mechanisms are in place to protect the enclave boundary, vulnerability checking and scanning procedures need to be implemented and exercised on the network and on the firewall.



**Figure 6.1-6 Case 1 Private to Public Network Communication**

As part of the Boundary Protection plan a site survey should be performed to ensure that the network operations and configuration is well understood. To assist with the site survey, a mapping tool can be used to construct the networks' topology and to examine the physical security of the network. The network map should detail which systems connect to public networks, and which addresses occur on each subnetwork. The network map should also identify which systems need to be protected from public access and identify which servers need to be visible on the outside and perimeter networks and what type of authentication and authorization is required before users can access the servers. The site survey should also examine which applications are used by authorized users of the network, what the anticipated growth of the network is, and what a users' privileges are including system administrators and firewall ad-



ministrators. In general, the site survey that should be attempted is directly related to the following.

- Technical expertise of the individual conducting the scanning.
- Level of threat.
- Sensitivity of potentially vulnerable information.
- Integrity of the source of the scanning software.

The placement of the firewall is of critical importance to the security of the network. The network needs to be configured to ensure that if an intruder accesses one part of the system, the intruder does not automatically have access to the rest of the system. A firewall should be placed at egress points to the network.

The recommended procedures that should be implemented relative to the firewall for protecting the enclave boundary include:

- Ensure that the virus scanning application is no more than a few weeks old. Viruses may infect the firewall itself as well as resources behind the firewall.
- Ensure that passwords and logins are not in clear text. Clear text passwords and logins are unencrypted and unscrambled and therefore vulnerable to sniffers on the Internet, allowing hackers to obtain passwords.
- Ensure that passwords and Secure Sockets Layers (SSL) are not cached by proxy agents on the firewall.
- Train personnel on firewall operations and administration.
- Audit for intrusive or anomalous behavior employing operating system, browser, and e-mail built-in audit capabilities.
- Routers can be configured as a firewall and for port mappings. With routers, anti-spoofing can be implemented, especially at the enclave boundaries or between domains of network administration. Source address spoofing denial-of-service protection can also be provided with access lists. The goal of creating an access list at the firewall level to prevent spoofing is to deny traffic that arrives on interfaces on nonviable paths from the supposed source address. For example, if traffic arrives on an interface sitting on the corporate side, yet the source address states that the traffic originated from the Internet, the traffic should be denied, as the source address has been falsified, or “spoofed.” Antispoofing access lists should always reject broadcast or multicast traffic.
- Routers could also be configured to hide the real network identity of internal systems from the outside network through port address translation. Port address translation minimizes the number of globally valid IP addresses required to support private or invalid internal addressing schemes.
- Configure operating system, browser, and applications for firewall functions and to permit specific access (make use of a proxy-based/application gateway). All traffic passing through the firewall should be proxied and/or filtered by the firewall. Proxies reduce the probability that flaws in the service can be exploited. Filtering limits the services that can be used and the user communities that have permission to use a service. The fewer services allowed through the firewall, the fewer opportunities there are to attack the protected network/system.

- Develop and exercise plans to handle any security incidents that may occur. These plans need to cover such things as:
  - How to handle detected port scans or more malicious attacks,
  - Recovery from any incident that degrades the performance of the network, and
  - The procedure for adding new services to the firewall.

## Case 2

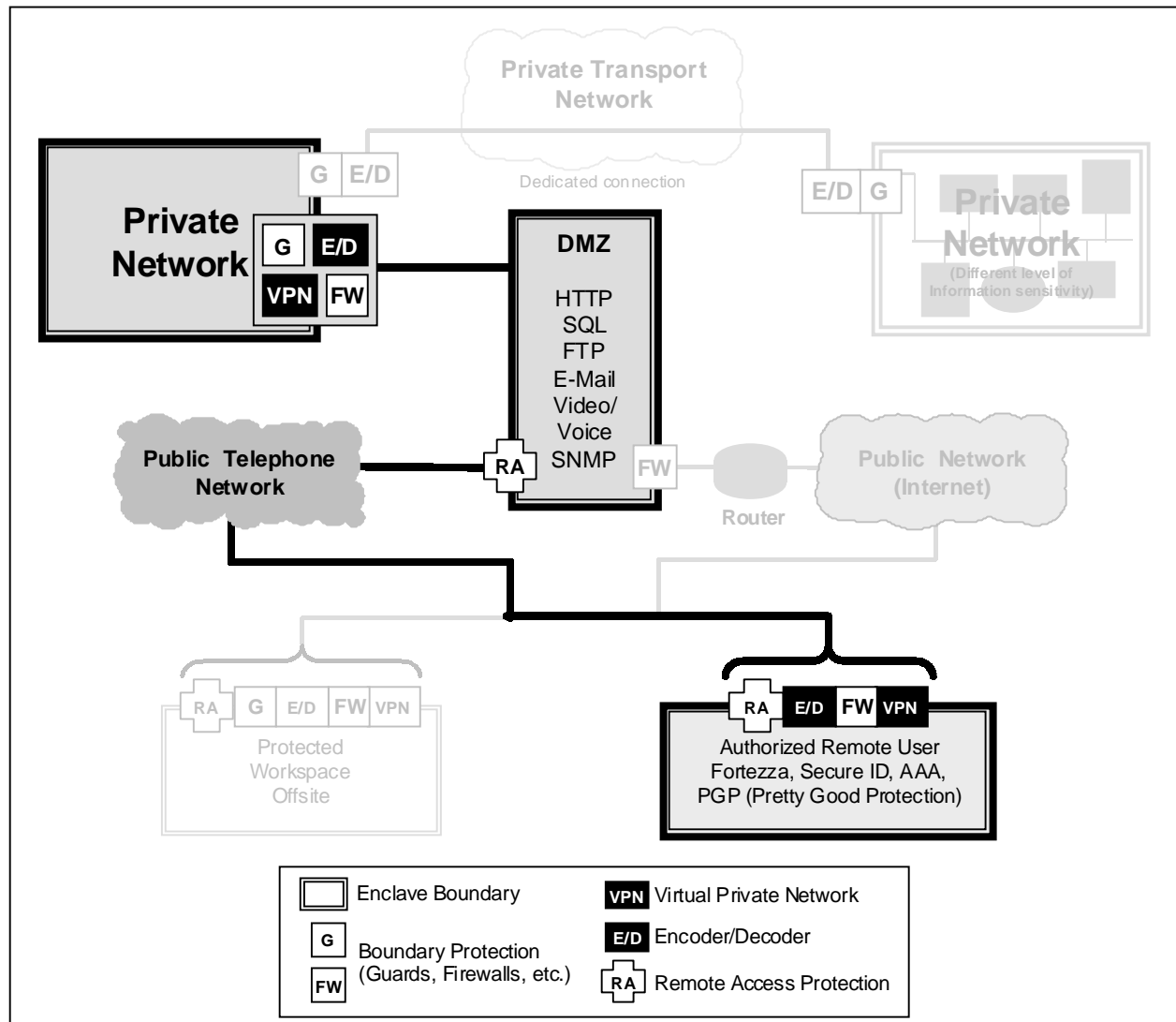
### ***A privileged user remotely connecting to a private network from dedicated workstations situated within a DMZ of a different protected network.***

This case is an example of remotely accessing a company's network from an off-site location. This off-site location is a protected network and has dedicated workstations connecting through that corporation's DMZ. Multiple connections through the DMZ can be established. Figure 6.1-7 illustrates a valid remote user connecting through the DMZ to the protected network. A DMZ allows authenticated authorized users to tunnel through the firewall. A DMZ also allows access to a Web or FTP server inside the firewall without exposing the rest of the network to unauthorized users. Otherwise, intruders could get control over the FTP or Web server and attack other hosts in the network. Therefore, servers should be placed so they can be accessed from any address in a separate subnetwork. Organizations can design, deploy, and proactively update and monitor a multizoned security network through a single firewall strategy. Administrators can create multiple DMZs within the network by simply adding rules to the existing firewall.

Modem banks should be established as part of the firewall protection approach so that users can dial out and remote users can dial in via a modem bank. Modems should not be allowed on a networked computers within the protected enclave boundary. By bypassing the implemented firewall and using a modem to connect to the Internet, all control over network security is lost. By using modem pools (a single dial-in point), all users are authenticated in the same manner. In addition, anti-spoofing controls can be applied at dial-up pools and other end-use connection points (also refer to <http://www.ietf.org/rfc/rfc2267.txt?number=2267>, RFC 2267). [5]

Before a user can access anything on the network, a username and password check should be completed. A stringent password policy is beneficial. One-time password schemes can also be used to further enhance the password security policy when establishing remote connections. Remote access connections use standard authentication techniques (please refer to Section 6.1.5, Firewall Technology Assessment, for more information regarding authentication).

Authentication, Authorization, and Accounting (AAA) for network access provides an additional level of security. AAA is the act of verifying a claimed identity, determining if the user has permission to access the requested resource, and collecting resource usage information for analyzing trends, auditing, billing or allocating costs. Message authentication plays a role when handling encrypted information. This verifies that the purported message sender is the person who really sent the message and that the message contents have not been altered. Although data can be authenticated at any hop on the way to the end destination, only the final destination may decrypt the data.



**Figure 6.1-7 Case 2: Remotely Accessing a Private Network**

Please refer to <http://www.ietf.org/internet-drafts/draft-ietf-aaa-na-reqts-07.txt>. [6] When remotely connecting to a company system, an alternative that also provides security is to establish a VPN. Please refer to Section 5.3, System High Interconnections and Virtual Private Networks.

Encryption of data is another common security measure. Encryption may be co-located with the firewall to provide secure tunnels to remote authorized users. Encoder/decoder products can be hardware- or software-based. Hardware-based solutions include PC cards (i.e., FORTEZZA), smart cards, or separate boxes attached to a network (for example, TACLANE, FASTLANE). For more information about FORTEZZA, please refer to <http://www.armadillo.huntsville.al.us/>. [7] There are also encryption software packages for encrypting e-mail such as Pretty Good Privacy (available free on the Internet, the site address is <http://www.wtvi.com/teks/pgp/>). [8] Software-based encoders/decoders also offer the capability of remote authentication, remote control, auto-answer secure data, and operation in both attended and unattended environments, therefore providing protection for facsimiles, e-mail,

and computer communications. For further information on the FASTLANE and TACLANE please refer to the FASTLANE category under Products & Services on General Dynamics' web page, [www.gd-cs.com](http://www.gd-cs.com). [9]

Users can also connect to their company's intranet via the Internet from a remote location. If a company's intranet is not configured properly, with some modification to the Internet site's URL, a hacker can gain access to the private intranet site. When setting up an intranet, access should be restricted to internally managed IP addresses only. Subnetting and access lists should also be implemented to allow only those permissible users within a company access to the Internet or certain intranet sites. Also, when establishing a virtual web or naming web pages, make the names cryptic so the content is not obvious and make all pages that contain private information password protected. This will prevent unauthorized people—from outside and inside the organization—from gaining unauthorized access to information.

## Case 3

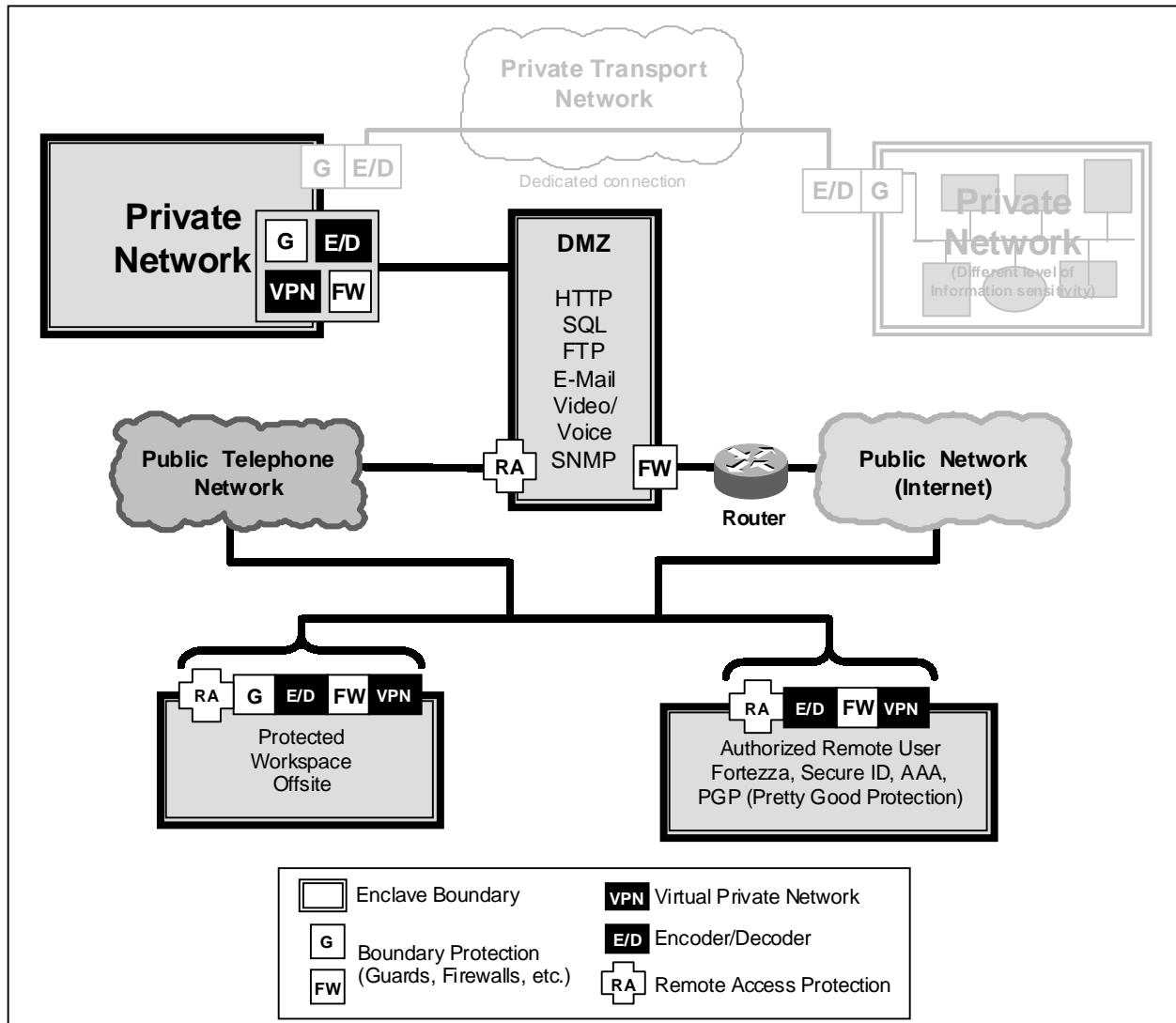
***Sensitive private network containing valuable information communicated through a lower level network to another network of equal classification / value (system high interconnects).***

This case involves networks that are interconnected at essentially the same information sensitivity level, using a lower sensitivity level unprotected, public transmission media (Internet, wireless). Referring to Figure 6.1-8, this scenario begins with the protected network containing proprietary data connecting via a public network to remote protected workspaces or valid remote users. At a minimum, this case requires:

- A boundary protection device (Firewall),
- A secure data connection device, i.e. encoder / decoder (KG, FASTLANE, TACLANE, FORTEZZA or other Commercial-Off-The-Shelf (COTS)/Government-Off-The-Shelf (GOTS)), and
- A proactive audit capability to include COTS/GOTS intrusion detection products.

Medium assurance levels are required for the enclave boundary protection implementations. For this case, the recommended boundary protection procedures that should be implemented in priority order are as follows.

- Institutionalize border security awareness and procedures. As outlined in Chapters 3 and 4.
- Configure the local computing environment (home network) with built-in features and services for enclave boundary protection. Installation of firewall and or comparable firewall feature set technology.
- Enable available audit capabilities. to include firewall ingress and egress points and auditing of attempted resource connections.
- Scan for viruses. Using current virus definitions and profiles. Ensure that definition file databases are no more than 3 months old.
- Perform a nonhostile vulnerability scan. Nonhostile scans include scans of: HTTP, FTP, Post Office Protocol (POP), SMTP, SNMP, ICMP, Telnet, Netbios, ensuring no deviations from initial network baseline scan.



**Figure 6.1-8 Case 3: Private Network Connectivity via a Lower-Level Network**

- Perform comprehensive vulnerability scans to include: scans for nonstandard UDP/TCP ports, unauthorized protocols, shares, nonencrypted passwords, potential operating system related vulnerabilities.
- Add intrusion detection. Intrusion detection methods should include the ability to proactively monitor packets, log and alert appropriate personnel based on level of threat/probe, identify and record addresses of threat initiator(s).
- Couple scanning, monitoring, and testing with intrusion detection. A network is only as strong as its' weakest link., By coupling scanning, monitoring, and testing—with intrusion detection—weaknesses and potential threats can be proactively identified upon first appearance or during the manifestation stage.

In addition, it is recommended that at least one staff person with an understanding of boundary protection be employed to configure and monitor the security parameters, perform virus and vulnerability scanning, and continually update the boundary protection and other security

measures as vulnerabilities are detected and new intrusion detection capabilities become available.

Software associated with the operating system, firewalls, and routers should be updated as the software continues to evolve with respect to built-in security features, especially as they relate to authentication and intrusion detection.

## Case 4

***Collaborating organizational LAN connecting to the main backbone network, which is of the same classification, with public WAN connections to remote protected networks, e.g., North Atlantic Treaty Organization (NATO) or foreign trusted network connected to main backbone network which is also connected to remote protected LAN(s) via a public WAN (Internet).***

This case involves connections that may jeopardize interconnected high-level systems if users and administrators are not aware of the public level WAN connection. As Figure 6.1-9 depicts, the unprotected network with proprietary data connects across a dedicated connection to the protected network with proprietary data, which is also connected to the public network/Internet and to remote users. The most basic level of protection for an enclave boundary includes employing the best available boundary protection technology (e.g., high assurance guards and intrusion detectors). Frequent virus and vulnerability scanning should also be performed by highly skilled personnel. An extensive security awareness program with institutionalized procedures for reporting and tracking is mandatory.

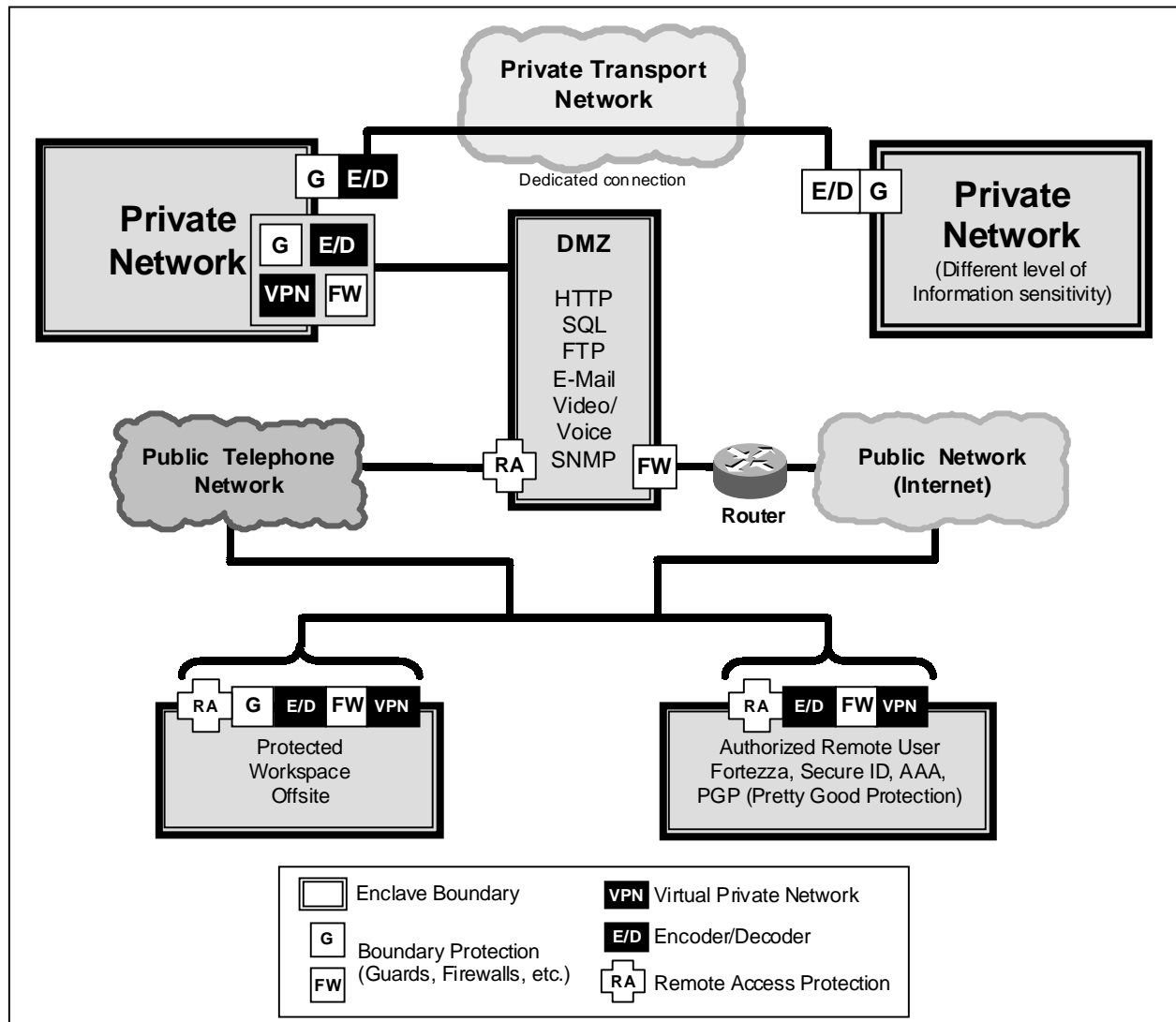
The following scenarios require comprehensive protection of enclave boundary penetrations or network access points, employing the best available technology.

***Collaborating LAN connecting to main LAN via dedicated connection.***

The collaborating LAN (foreign company, NATO agency, etc.) is of the same information sensitivity level, and the anticipated threat level is at a minimum. Because the collaborating agency is accessing peripheral data, limited network resource access is required. Full access to all enclave contained information assets is not needed. Initiating an internal proxy server with a strict access security list is recommended (protected Solaris, local/global user access list via NT's NT File System (NTFS) with auditing enabled, etc.). The collaborating LAN should be connected via a secure means, either through a data encoder/decoder (KG) or similarly approved security device. Intrusion detection monitoring products should include real-time auditing and tracking capabilities.

***Protected off-site LAN with same security level connecting to main LAN via public WAN (Internet) with main site having a directly connected collaborating site.***

All previously outlined security precautions need to be met (as defined by case studies 1, 2, and 3). The main LAN needs to have a strict access list in place (protected Solaris, local/global user access list via NT's NTFS with auditing enabled). This precaution is to ensure that the connected collaborating LAN is able to access only predetermined enclave information assets, including resources at the main LAN as well as the offsite protected resources. To further ensure that only approved data is exchanged from the off-site LAN to the collaborating agency, it is recommended that guards be installed at both the ingress and egress location on the enclave boundary of the home enclave LAN.



**Figure 6.1-9 Case 4: Collaborative LAN's with Public Network Connections**

The guards are present to ensure that only approved filtered data is exchanged between trusting and trusted networks/domains. Implemented intrusion detection monitoring products need to include real-time auditing and tracking capabilities.

*Collaborating LAN connecting to protected remote site using main LAN's backbone.*

All previously outlined security precautions need to be met (as defined by case studies 1, 2, and 3). If the collaborating LAN needs to connect directly to the off-site LAN without accessing any main LAN resources the following need to be addressed.

- A router or layer 3 switch is needed at the point of presence of the main LAN.
- A static route needs to be configured to route traffic directly to the off-site LAN, via the main LAN's backbone.
- Data traffic needs to travel over the main LAN's encoders/decoders and through its DMZ.

- A guard needs to be installed at the boundary of the off-site LAN.

The purpose of this type of configuration is to prevent a direct association between an off-site and collaborative LAN (i.e., a foreign organization/agency that is communicating with a local company or agency, the main LAN, acts as a go-between).

For this case and the associated scenarios, the recommended boundary protection procedures are similar to the previous recommendations, but require higher assurance boundary protection technology implementations. The following recommendations should be implemented as a comprehensive package with reference to which scenario the network most resembles.

- Institutionalize boundary security awareness and procedures. As outlined in Chapters 3 and 4.
- Configure the home enclave network using built-in features and services for boundary protection. Installation of firewall and or comparable firewall feature set technology.
- Enable available audit capabilities. To include firewalls, ingress and egress points and auditing of attempted resource connections.
- Scan for viruses. Using current virus definitions and profiles. Ensure that definition file databases are no more than 3 months old.
- Perform a nonhostile vulnerability scan. Nonhostile scans include scans of: HTTP, FTP, POP, SMTP, SNMP, ICMP, Telnet, Netbios, ensuring no deviations from initial network baseline scan.
- Frequently perform comprehensive vulnerability scans; include: scans for nonstandard UDP/TCP ports, unauthorized protocols, shares, nonencrypted passwords, potential operating system related vulnerabilities.
- Incorporate enterprise-wide intrusion detection. Intrusion detection methods should include the ability to proactively monitor packets, log and alert appropriate personnel based on level of threat/probe, identify and record routing addresses of threat initiator(s).
- Incorporate infrastructure attack “early warning.”
- Employ supplementary boundary protection between off-site locations. (firewall/guard services).
- Couple scanning, monitoring, testing, and intrusion detection. A network is only as strong as its weakest link. By coupling scanning, monitoring, testing, and intrusion detection, weaknesses and potential threats can be proactively identified upon first appearance or during the manifestation stage.

## 6.1.7 Enclave Boundary Protection Framework Guidance

The technologies discussed in this section and the types of techniques they employ should typically be composed to form a solution set to defend the enclave boundary. Although the technologies overlap, each focuses on a different subset of security countermeasures. Additional access control mechanisms should also be used in forming mitigation approach sets. These in-



clude encryption or application layer discretionary access controls to permit or deny access to specific data within an enclave. Given these countermeasures, it must be determined how, where, in how many places, and how many times they should be applied. Places to which the countermeasures can be applied include at the enclave boundary, workstation/LAN interface, individual workstations, servers, operating systems, or at the application level. A layered security approach can be used, determining how many places a countermeasure should be applied. How many times a countermeasure should be applied is the choice between per session authentication and per packet authentication. It must also be determined how strong the security measures must be.

A number of factors generally influence the selection of firewall approaches. The mission needs and services desired by the users are primary factors in shaping mitigation approach sets. The risks to a given system must be assessed in terms of:

- The differences in information value and threat between the protected enclave information assets and the external networks to which it is connected,
- The environments and architecture, and
- The impacts of potential attacks.

In addition, cost, policy mandates, scalability, maintainability, and overhead (including performance degradation and manpower) must be considered. Clearly the specific protection approaches and products selected must also be those that can address the specific services, protocols, operating systems, applications, components, etc. employed in the user's environment. Ideally, the technologies which incorporate all prescribed countermeasures, at the appropriate levels, and addressing all aspects of the specific user environment should be implemented. As indicated in Section 6.1.5, Firewall Technology Assessment, and below there are gaps in successful achievement of countermeasures, performance, and other areas.

Potential negative impacts are associated with any of the technology solutions. Desired performance of a firewall must be determined when implementing a firewall to defend the enclave boundary. There is a tradeoff between speed and security. A network can be more secure when the firewall performs more checking on the packets. However, the amount of checking that a firewall performs has an effect on the volume and the speed at which traffic can transverse the enclave boundary protection.

In addition, while greater restrictions to operations does yield greater protection of the enclave assets, the restriction of dangerous operations also restricts useful operations. There comes a point at which the tradeoff for greater security becomes more than the users want to pay in lost capability or hampered performance. For example, some antiviral and disinfectant (subversion-constrained) software may actually do as much damage to operational performance as viruses themselves might. Some systems may fail to prevent infections but prevent the user from getting rid of the virus. Some antiviral systems may actually delete files without alerting the user or offering alternative approaches. Disinfecting has been known to leave workstations in a worse state than the infection did. The primary approach to selection of a security protections should be to maximize benefits while minimizing harm. Only through a comprehensive risk analysis, with knowledge of the characteristics and tradeoffs of different technologies and specific products including cost and resource constraints, can effective enclave boundary protection be implemented and maintained.

The first step in any effort to implement an enclave boundary protection mechanism and additional technology to protect the enclave information assets is to develop a security policy. The

boundary protection mechanisms will then serve to implement this security policy. An in-depth requirement analysis forms the basis for the development of the policy and subsequent selection of a protection devices.

Clearly, the environment in question will dictate the level of security robustness. For example, in connecting enclaves of different classifications, whether through a direct connection or through another network, additional security precautions must be taken. Remote access to the enclave through the boundary protection mechanism will require security mechanisms designed specifically for this situation. Firewalls, for example, generally have the capability to form an encrypted link to the remote user. Boundary protection mechanisms, which are used inside the enclave to limit access to restricted information, on the other hand, tend to be cheaper and less complex than those devices located at the boundary of the entire enterprise. Firewall technology has evolved so that firewalls are now developed and marketed specifically for intranet firewall applications.

In addition to the specific environment in question, there are a number of general tradeoffs, which should be addressed when implementing firewall technology. One important tradeoff with regard to firewall technology is between security and ease-of-use. The more rigorous the checks for user identity and user activity, the more inconvenience the user must endure. On the other hand, if the firewall simply passes everything through to the internal network, security is inadequate, even for the least sensitive data. In choosing a firewall, both the needs of the users for services and the security requirements must be balanced; otherwise, the users will find ways to bypass the firewall, weakening the protection of the enclave boundary.

Packet filters and stateful packet inspection technologies focus on flexibility. In general, these firewalls are able to support many services, and additional services can be easily added. However, this flexibility comes with a price. It is quite easy to configure these types of firewalls to permit dangerous access to services through the firewall. The ease-of-use administrative interfaces and preconfigured support for many services lend themselves to configuration errors. Application gateways, on the other hand, provide better auditing and finer grained control. For example, application gateways can be used to allow certain activities, such as sending a file to an untrusted network, while blocking a user from copying a file from an untrusted network. In general, router-based firewalls are best for a dynamic environment where lots of things change in a short time frame. Application-level firewalls are better if a more deliberate approach to security is necessary.

Other considerations in selecting a firewall include the skill level available for maintaining the firewall. As noted above, proper configuration and maintenance of the firewall is a critical security element. If an organization does not have the staffing to assign qualified personnel to operate and maintain the firewall, there are options to purchase firewall maintenance services, from either the firewall company or the ISP. These costs of staffing or services should be considered, as well as the corporate credentials of the firewall vendor, and the quality of the documentation available with the firewall.

# References

1. B. Frasier. Site Security Handbook RFC 2196. September 1997  
<<http://www.ietf.org/rfc/rfc2196.txt?number=2196>>.
2. SOCKS. 1 May 2000 <<http://www.socks.nec.com>>.
3. FTP Directory. 1 May 2000 <<ftp://ftp.nec.com/pub/socks>>.
4. Rekhter Y., et al. "Address Allocation for Private Internets. RFC 1918." February 1996  
<<http://www.ietf.org/rfc/rfc1918.txt?number=1918>>.
5. Ferguson P. and D. Senie. "Network Ingress Filtering: Defeating Denial of Service Attacks which employ IP Source Address Spoofing." 18 May 2000  
<<http://www.ietf.org/rfc/rfc2267.txt?number=2267>>.
6. AAA Working Group. "Criteria for Evaluating AAA Protocols for Network Access". 26 April 2000. On line posting. 11 May 2000  
<<http://www.ietf.org/internet-drafts/draft-ietf-aaa-na-reqts-05.txt>>.
7. FORTEZZA Cryptography of the 21<sup>st</sup> Century. 12 May 2000.  
<<http://www.armadillo.huntsville.al.us/>>.
8. Pretty Good Privacy Software. 12 May 2000 <<http://www.wtvi.com/teks/pgp/>>.
9. General Dynamics Communications System. 12 May 2000 <[www.gd-cs.com](http://www.gd-cs.com)>.

# Additional References

- a. Cisco Systems, Inc.. "How Data Moves Through The Firewall." 19 May 2000  
<[http://www.cisco.com/univercd/cc/td/doc/product/iaabu/pix/pix\\_v41/pixcfg41/pix41int.htm#xtocid297201](http://www.cisco.com/univercd/cc/td/doc/product/iaabu/pix/pix_v41/pixcfg41/pix41int.htm#xtocid297201)>.
- b. Computer Security Resource Center. 1 May 2000 <<http://csrc.nist.gov/>>.
- c. Internet/Network Security. 1 May 2000  
<<http://www.netsecurity.about.com/compute/netsecurity>>.
- d. Defense Information Systems Agency. Firewall Configuration Guide, 12 June 1998.
- e. Internet/Network Security site. "The Secure Telecommuters FAQ" Page 10 May 2000  
<<http://netsecurity.about.com/compute/netsecurity/library/weekly/aa020200c.htm>>.
- f. National Security Agency/V43. Department of Defense Firewall Guidance. Version 1.0 Draft, 31 March 2000.
- g. Network Vulnerability Analysis and Penetration Testing. 8 May 2000  
<<http://www.blackmagic.com/assessment.html>>.
- h. The Source of JAVA™ Technology. "Applets." 8 May 2000  
<<http://www.java.sun.com/applets/index.html>>.

- i. United States Navy Web Information Service. 12 May 2000  
<<http://infosec.navy.mil/products/securevoice/stu3.html>>.  
Enter at <<http://infosec.navy.mil>>, then, navigate to:  
< <http://infosec.navy.mil/products/securevoice/stu3.html>>.

## 6.2 Remote Access

Remote access enables traveling or telecommuting users to securely access their Local Area Network (LANs), local enclaves, or local enterprise-computing environments via telephone or commercial data networks. Remote access capability draws on both the Virtual Private Networks (VPNs) and the Defending the Enclave Boundary sections of this document. The remote access user connects by a shared commercial path, and can maintain the privacy of their connection using encrypting modems, technologies applicable to Vans needs (as discussed in Section 5.3 (System High Interconnections and Virtual Private Networks (Vans)), or other technologies suitable to this requirement. Because the user entry point into the enterprise-computing environment could be used by a hostile connection, the enterprise must implement enclave boundary protection (as discussed in Section 6.1, Firewalls). The remote user's computing assets are also physically vulnerable, requiring additional protection. This section draws on the preceding two and explores protection for information storage to address the specific problem of remote access.

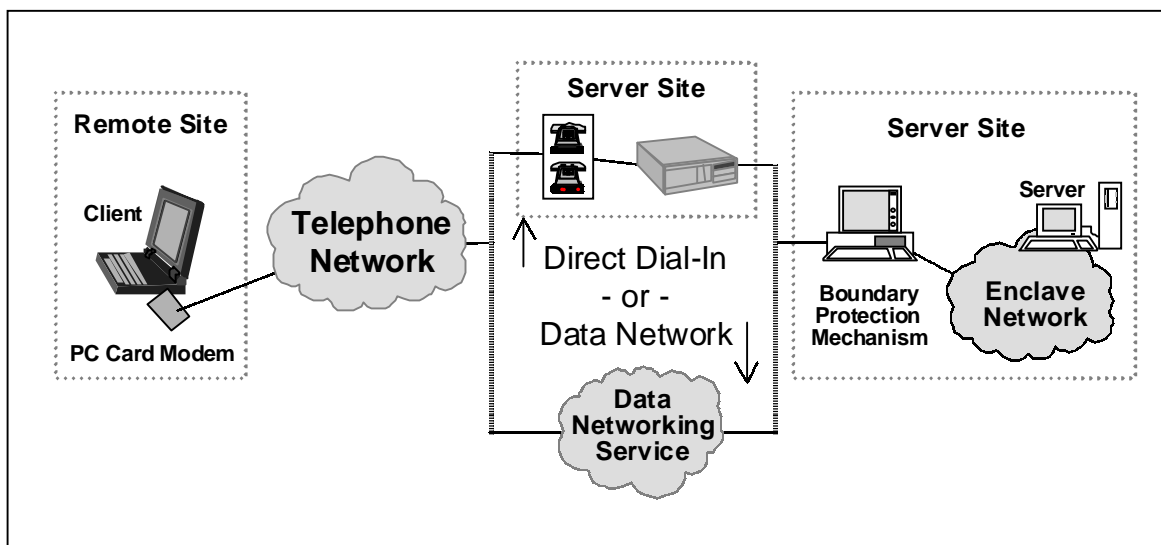
### 6.2.1 Target Environment

Within this section, traveling users and telecommuters are both treated as remote users. However, the environment of these two groups differs in the degree of physical exposure of the remote computer. The traveler's computer is vulnerable to theft and tampering while the user is in transit and while their computer is in storage. These risks are particularly great overseas. The telecommuter's computer is also vulnerable to theft and tampering, but to a much lesser extent if the physical location of the hardware is within Continental United States (CONUS). In addition, because the telecommuter's remote location is relatively fixed, additional steps can be taken for physical protection that are not feasible for traveling users. Conversely, the telecommuter's fixed remote location makes targeting by an adversary easier than in the case of mobile traveling users.

As depicted in Figure 6.2-1, remote users access their enterprise-computing environments by communication paths shared with others. Many remote users employ the Public Switched Telephone Network (PSTN) to access their home enclave directly or use the PSTN to connect to a data network such as that of an Internet Service Provider (ISP), which in turn connects users to their enterprise-computing environment. Other remote users employ emerging network technologies, including digital wireless service, cable modems, Integrated Services Digital Network (ISDN), and other high-data-rate media. Remote access via these networks increases the level of threat and imposes architectural constraints to the security solution. This section of the Information Assurance Technical Framework (IATF) treats remote access, via these networks, separately from direct dial-in to an enterprise-computing environment via PSTN.

Note that for this section, remote access is limited to the capability of providing access to the information contained in users' home system-high LANs, enclaves, or enterprise-computing environments from remote locations, which—during the period of connectivity—are assumed to be controlled at the same system-high level as the home system. In other words, remote users with authorized access to Sensitive-But-Unclassified (SBU) information will be given ac-

cess to SBU information contained in their home SBU system-high enclaves and remote users authorized access to Secret information will be given access to Secret information contained in their home Secret system-high enclaves.



**Figure 6.2-1 Typical Remote Access Environment**

In the case of Secret remote connectivity, the proposed remote connectivity approach will give the remote user the ability to store information on the remote terminal (typically a notebook computer) hard drive in an encrypted format, thereby declassifying the terminal when it is not in operation. However, during the period of connectivity to the home system, the remote user must provide sufficient physical protection and safeguarding of the Secret information being processed.

## 6.2.2 Consolidated Requirements

### 6.2.2.1 Functional Requirements

The following requirements are from the user's perspective.

- Remote users should have access to all information stored on their remote computers, stored on their home enclave workstation, or available within their home enclave information infrastructure. Because remote users need to conduct their business using familiar tools while traveling to a remote location, cryptographic application interfaces on the remote user's terminal should be similar and have the "same look and feel" as those provided at their home enclave. Applications which may be launched from a system-high enclave as a result of a remote user request, shall continue to support all security services as required by the enclave system security policy and procedures.
- The user should know when security features are enabled. Indications should not be intrusive, but the user should be able to tell easily when security features are working,

and more important, when they are not. Feedback to the user is very important in any security solution.

- The security solution should have minimal operational impact on the user. It should not impose a significant performance penalty, or require extensive training.
- The traveling user's security suite should not include any external devices. Some remote users simply do not have room for these devices in their computing packages. Solutions that are unobtrusive to the user (e.g., user tokens and software products) are preferred.
- The remote user's equipment should be unclassified when it is unattended. Both the data stored on the remote user's computer and the approved configuration of the remote user's computer must be protected from unauthorized disclosure, modification, or manipulation when out of the direct control of the authorized remote user. This protection must effectively protect the computer and stored data from compromise if the computer is lost, stolen, or used to communicate with lesser security level authorized hosts. Assuming the data stored on the remote user's equipment is appropriately protected, the user is required to safeguard the terminal as would be required of high-value items.
- The remote user should not have greater access than would be available if accessing the enclave information resources from within the enclave.

### 6.2.2.2 Interoperability

Remote access systems that implement interoperable solutions facilitate the movement of users between organizations and increase the likelihood that the system can be supported and upgraded in the future. Interoperability also provides for the maximum evolution of this security solution in the commercial marketplace. For these reasons, the following interoperability requirement is added.

Security solutions should be based on open standards. The use of proprietary implementations creates significant issues related to interoperability and logistics support. To ensure an effective solution, the remote access mechanism should integrate easily into existing information systems and provide a path for upgrading to emerging technology (as discussed below).

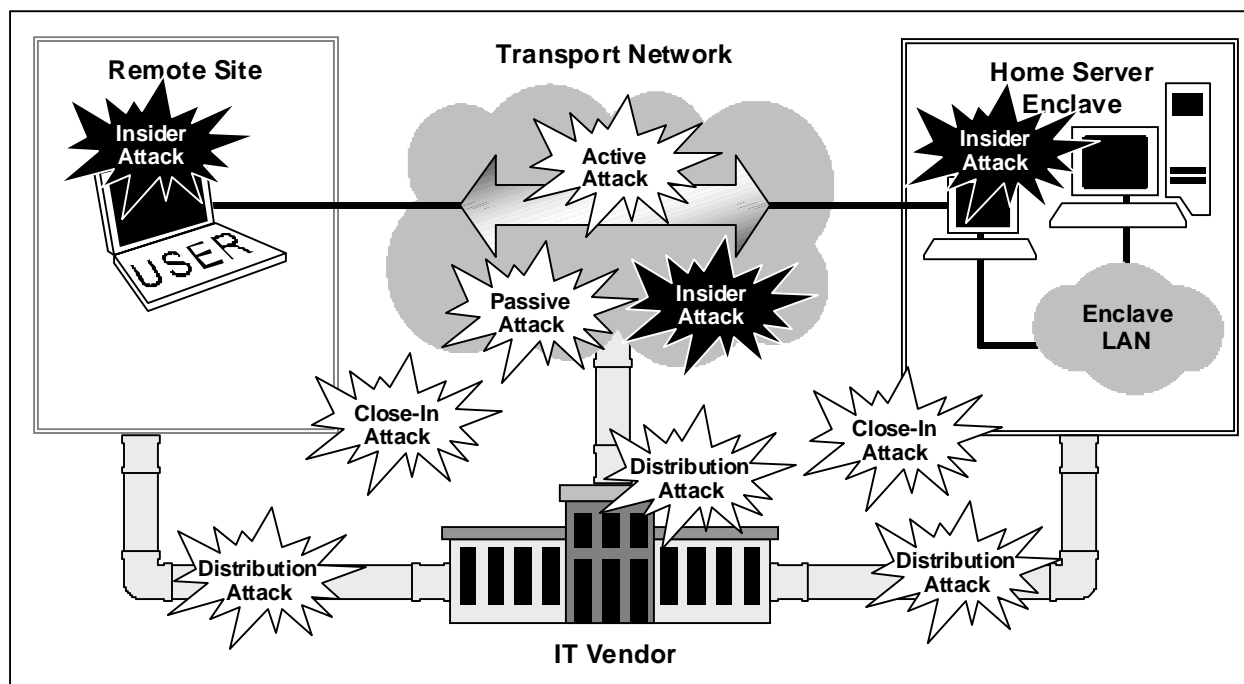
### 6.2.2.3 Emerging Technology

It is desirable that the security solutions be capable of evolving to higher data rates and be adaptable to alternative means of communication, such as cellular telephony, wireless networks and ISDN.

## 6.2.3 Potential Attacks

All five classes of attacks introduced in Chapter 4 Technical Security Countermeasures are of concern in the remote access scenario. Section 6.1, the Firewalls section goes into detail on network attacks. The VPN's section's (Section 5.3) treatment of passive, network and insider at-

tacks is directly relevant to remote access. Since proper configuration and execution of software is critical to the proper functioning of security mechanisms, distribution attacks are also a concern. Remote access places the user's computer in public environments, adding the possibility of physical attack to the five generic attack classes. With reference to Figure 6.2-2, the following summarizes potential attacks against the remote access scenario.



**Figure 6.2-2 Attacks Against the Remote Access Scenario**

### 6.2.3.1 Passive Attacks

An attacker monitoring the network could capture user or enclave data, resulting in compromise of information. Capture of authentication data could enable an attacker to launch a subsequent network attack. Analysis of traffic captured by passive monitoring can give an adversary some indication of current or impending actions. Compromising emanations could also be intercepted.

### 6.2.3.2 Active Attacks

These attacks are most likely to originate from the Internet, but, with more effort, could also be mounted through the PSTN. Also attacks can target the remote user's computer, the user's enclave, or the user's connection to the enclave, potentially resulting in the loss of data integrity and confidentiality, and ultimately in the loss of use of the network by authorized users (e.g., a denial-of-service attack).



### 6.2.3.3 Insider Attacks

An insider is anyone having physical access to the remote user's computer or the network enclave from within the user organization's corporate boundaries. These attacks could be motivated by malice or could result from unintentional mistakes by the user. Deliberate attacks can be especially damaging to the organization's information system due to the attacker's access to the information, their advantage of knowing the network's configuration, and thus their capability to exploit the network's vulnerabilities.

### 6.2.3.4 Distribution Attacks

Distribution attacks could occur at the Information Technology (IT) provider's site while the product is developed, manufactured and shipped, while the remote user's computer is being configured, or when software is passed to the user's computer (including software passed over the network). This type of attack could result in a network's device (e.g., firewall, router, etc.) being used to perform a function for which it was not intended, thus making the remote access capability or the enclave vulnerable to attack.

### 6.2.3.5 Close-In Attacks

The remote user's computer is subject to theft and tampering. Physical attack could also result in theft of the traveling user's computer, a denial-of-service attack. Typically, there are non-technical countermeasures (e.g., procedures) available for dealing with physical threats. The Framework addresses these since these are also technical countermeasures available that could help to mitigate those threats.

## 6.2.4 Potential Countermeasures

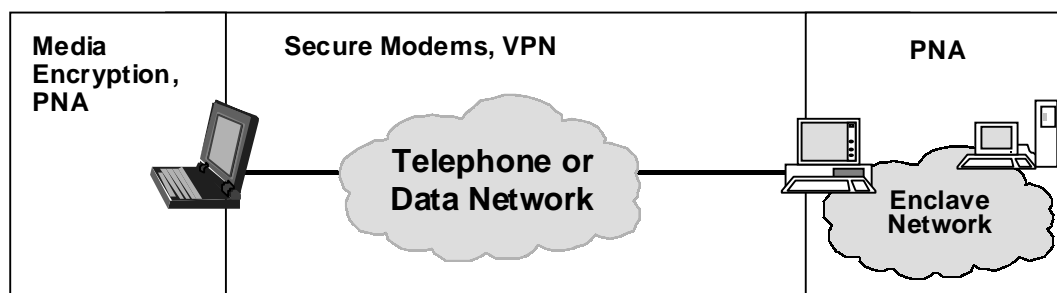
The following security services are required to counter the potential attacks against the enclave.

- Strong and continuous user authentication should be the basis for allowing access to the enclave. Strong continuous two-way authentication protects the enclave, the remote user and the connection from network attacks. Cryptography-based authentication at the enclave boundary ensures that only authorized users can gain access to the network. Use of a boundary protection mechanism is used in conjunction with cryptography-based authentication to provide a basis for controlling a user's access to individual network services. Continuous authentication prevents an unauthorized user from hijacking the remote user's session.
- Confidentiality may be invoked for all information flowing between the enclave and the remote user's computer. Confidentiality guards the enclave and the remote user from passive intercept attacks. Although, encryption does little to guard against traffic analysis, the data and metadata (information about data) are protected against direct intercept and compromise. This security service is dependent, of course, on the level of required protection afforded the data.

- The information in the remote user's computer should be protected:
  - 1) When the computer is not in use. This protects the information in case of theft of the workstation, or unauthorized physical access.
  - 2) When the computer is connected to unclassified or untrusted networks. This guards against network attacks (e.g., session hijacking) from an unclassified and/or unauthorized network.
- The integrity of the remote user's hardware and software should be protected. Detection and protection mechanisms can guard against distribution attacks, tampering by an outsider, and physical access by an unauthorized user.
- The integrity of data flowing between the remote user's computer and his enterprise-networking environment should be protected. This protection is typically provided at the applications layer. See Section 7.1 Security for System Applications of the Framework for details.

## 6.2.5 Technology Assessment

The three technologies—media and file protection, workstation integrity, and enclave and connection protection—are included in this section and depicted in Figure 6.2-3 counters specific types of attacks. Some attacks, such as tampering, are only partially addressed by technical measures. Non-technical security measures, as discussed in Chapter 4, Technical Principles—physical protection of the laptop, prevention of casual “over-the-shoulder” observation of classified information—are critical to overall system security and should be considered a vital part of a remote access user policy. This section of the Framework only covers those technical measures that will counter attacks relevant to the remote access category.



**Figure 6.2-3 Security Technologies in the Remote Access Scenario**

### 6.2.5.1 Media and File Encryptors

In some cases, physical removal of the remote computer storage media (typically a hard drive) between remote connection sessions is not acceptable. Encryption of the information on the storage media can provide confidentiality and integrity, alleviating the need for physically removal of the media. Media encryptors and file encryptors protect the information in the computer in the event of unauthorized physical access to the computer. File encryptors can

protect the confidentiality and integrity of individual files, provide a means of authenticating a file's source, and allow the exchange of encrypted files between computers. Media encryptors protect the confidentiality and integrity of the contents of data storage media. For example, they can help maintain the integrity of the remote user's computer by verifying the Basic Input/Output System (BIOS) and ensuring that configuration and program files are not modified.

With the exception of some system files, media encryptors encrypt the entire contents of the drive. The media encryptors must leave some system files unencrypted so that the computer can boot from the hard drive. The integrity of most of these unencrypted system files can be protected by a cryptographic checksum; this protection will not prevent a tamper attack, but it will alert the user that that data has been altered. System files contain data that changes when the computer is booted cannot be protected.

File encryptors typically implement a Graphical Users Interface (GUI) that allows users to choose files to be encrypted or decrypted. This protects individual files, but it does not protect all files on the drive. Many applications generate temporary files that may contain user data. These files are normally closed (but not necessarily erased) when the application is terminated. However, the application does not terminate in an orderly fashion, these temporary files may remain open. Some operating systems do not actually erase data when files are closed or deleted. Instead, they alter the name of the file in the file allocation table or deallocate the storage locations on the media. The user's data then remains on the hard drive until the space is allocated to another file and overwritten. Thus, unencrypted and potentially classified user data can remain on the hard drive after system shutdown, either because of the application's failure to erase temporary files or by the design of the operating system's file closure function. For these reasons, media encryptors provide better protection for the information on the disk drive—especially while the computer is not in use—than do file encryptors.

Media encryption's robustness is an advantage only when proper key management is used in protecting the information. There must be provisions to allow trusted key management to protect the key when encrypting the media and when the key is in storage. See Section 6.2.7 (Framework Guidance) of this chapter for further discussion of the Secret dial-in case. Media encryption also supports workstation integrity, the topic of the next section.

## 6.2.5.2 Workstation Integrity

Workstation integrity components are necessary to protect the integrity of a remote computer's operation and data against active (network based) and software-distribution threats. Active attacks include attempts to steal data by circumventing or breaking security features or by introducing malicious code. The software distribution threat refers to the potential for malicious modification of software between the time it is produced by a developer and its installation and use on the remote user's computer.

Workstation integrity mechanisms to counter active attacks are addressed in the Firewalls section of the Framework. Products for detecting and removing computer viruses are available for both the workstation and boundary protection mechanism. Media encryption protects the configuration and software of the remote user's computer against malicious modification during the operational phase; it does not address this modification during the developmental or the distribution phases. Trusted operating systems can ensure the policy-enforced relationships

between subjects and objects, thus limiting any effects the malicious code introduced into the machine might have on the system's integrity.

Software distribution attacks are discussed in Chapter 4 (Technical Principles) (System Security Methodology). Most software distribution attacks can be thwarted by the use of digital signatures. Software can be signed at the manufacturer before distribution; these signatures are verified before the software is installed on the user's computer. Commercial file encryption packages containing this capability are available.

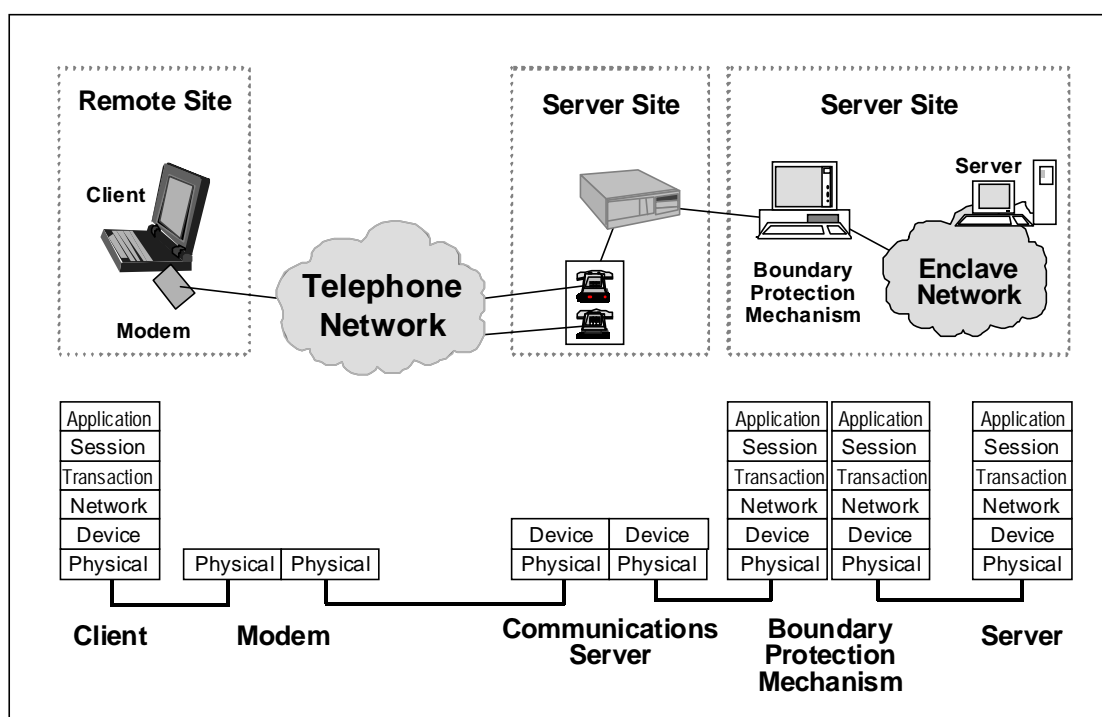
### 6.2.5.3 Enclave Boundary and Connection Protection

Components to implement authentication, confidentiality, and integrity mechanisms can operate at several layers in the protocol stack, with trade-offs in assurance, performance, and networks supported. Starting toward the bottom of the protocol stack, options include secure modems, data link layer technologies, network layer products, transport and session layer products, and application layer products. The protocol layer chosen does not necessarily imply a certain level of information assurance. There are mechanisms that can provide either at a high level of assurance, a low level of assurance, or something in-between at any protocol layer. Connection protection is dependent on an organization's risk management decision concerning the level of assurance placed on these mechanisms. All of these approaches, except application layer protocols are discussed in the VPN section (Section 5.3 System High Interconnections and Virtual Private Networks (VPNs)). The authentication mechanism should provide mutual authentication of the remote user and the enclave's boundary protection mechanism, which is described in the Firewalls and Guards sections (Sections 6.1 and 6.3, respectively) and shown in Figure 6.2-1. It also shows both options for connecting to the enclave—by direct dial-in to the enclave and by an ISP. Figure 6.2-4 shows the protocol layers associated with the remote access scenario.

## Secure Modems

Secure modems offer an inherent means of boundary protection: the identity of the remote user's modem is established by strong authentication before any network connections are initialized, preventing unauthorized modems from attempting an active attack. The invocation of encryption within a modem provides a high level of assurance provided that the encryption function is properly invoked and is protected from tampering. However, the implementation of additional features, such as plaintext bypass, can reduce some of that assurance. For instance, a secure modem needs a means of bypassing the encryption engine if it is also to interoperate with a nonsecure modem. Any bypass feature in a secure modem must be carefully implemented so that it is not possible to bypass the cryptography accidentally or maliciously.

Strong authentication requires a significant cryptographic processing capability both in the calculations required to validate a signature and in the verification of the identity contained in a certificate (e.g., checking against a list of authorized users). The identity that is established by modem authentication may not necessarily be made available to the network. This requires the remote user to log into the network separately.



**Figure 6.2-4 Protocol Layers In Remote Access Scenario**

## Data Link Mechanisms

Data link layer protocols such as Point-to-Point Protocol (PPP) and Serial Line Internet Protocol (SLIP) encapsulate network layer packets for transmission via modems. Security services can be applied to these protocols to allow authentication and protect the connection between the remote user and the home enclave's communication server. Unlike the large bandwidth data links discussed in the VPN section, the remote user's data link is dedicated, so authentication of individual users is possible. This assumes, of course, that the remote machine is dedicated to one (and only one) user since authentication at the data link layer relies on lower level physical addresses vice on higher layers that can distinguish among multiple users (e.g., with user Identifications (IDs)).

Data link mechanisms allow users to choose their own modem hardware and upgrade or change it at their convenience, provided that the hardware can interoperate with the enclave's boundary communications hardware. A server implementing a data link mechanism could use the results of cryptographic authentication as a basis for access to the enclave. Data link security mechanisms are likely to be implemented in workstation software, where processing power and memory are more readily available than in the case of special-purpose security hardware. This makes implementation functions such as continuous authentication and certificate path validation more practical. However, it also makes these functions dependent on the integrity of the workstation on which they are running and more vulnerable to implementation errors and subversion.

At the data link layer, no information is available about the network resources or services the remote user is attempting to access. Any filtering mechanism would need to be implemented at a higher layer of the protocol stack.

## Network Layer Mechanisms

Network layer protocols, such as Internet Protocol (IP), assign addresses to devices and pass data packets between them. ISPs assign an IP address to the remote user and pass IP packets for the remote user. For this reason, the network layer is the lowest layer at which security services can be applied in the ISP case. The VPN section addresses IP connections across public networks, and recommends the use of Internet Security (IPSec) with both Encapsulated Security Protocol (ESP) and Authentication Headers (AH). The VPN section also recommends the use of external encryptors. The current generation of external encryptors must be configured by a trained operator and are expensive and relatively bulky, so external encryptors are currently infeasible for remote access. However, IPSec mechanisms are implemented in network card hardware, in modem cards, and in software on the user's computer (as before, the proper functioning of software mechanisms depends on the integrity of the user's computer).

Network layer mechanisms allow strong authentication directly from the remote user's computer to the boundary protection device, allowing the boundary protection device to base access control decisions on the user's identity. Network layer information allows the boundary protection mechanism to filter access to individual machines in the enclave. The downside is that they leave all of the enclave's dial-in equipment before the network device—specifically the modems and the communications server—exposed to network attacks. Provided that the communications servers properly configured and controlled, the potential for successful attacks against a communications server is relatively low (except for denial-of-service attacks). Remote control and administration of these devices can make the network vulnerable to attack by providing potential access to root level privileges. Please refer to Section 6.1 (Firewalls) for more information.

## Transport and Session Layer Mechanisms

The transport layer forms a reliable channel between devices. The session layer establishes and synchronizes a communication session between two devices. The transport or socket layer is the lowest layer with information on the service being accessed so that security services can be called on a per application basis. The transport and session layers are discussed in the VPN section (Section 5.3). For the remote access scenario, layers share many of the advantages and disadvantages of network layer mechanisms—they can allow continuous authentication directly to the boundary protection mechanism and allow further access control decisions based on the cryptographically authenticated identity. Transport and session layer mechanisms are not likely to be hardware based, making them vulnerable to tampering and dependent on the integrity of the user's computer.

The Secure Socket Layer (SSL) protocol currently exists as Request For Comments (RFC) 2246. The Internet Engineering Task Force (IETF) is developing this protocol into a future standard, the Transport Layer Security (TLS) protocol, which sits at the top of the transport layer. Product implementations of socket mechanisms should comply with the IETF standard, which is currently SSL.

The Remote Access Dial-in User Service (RADIUS) protocol (RFC 2138) was designed to authenticate remote users using a shared secret. The RADIUS protocol is currently an Internet Draft published by the IETF. Authentication requests are handled by a centrally located authentication server, which provides a method of supporting the management of remote users. The access requests made by RADIUS clients are capable of carrying attributes that include user name, user password, client identification, physical port identification, or other information. When passwords are present, they are protected by using RSA MD5. The ability of RADIUS to support a wide range of client attributes used in access control decisions makes this protocol very flexible. Access privileges can be varied for each user, as well as for the access method each user attempts. Maintaining a central RADIUS server, which controls the privileges for each user, makes RADIUS authentication scalable to handle large numbers of remote users.

## Application Layer Mechanisms

Application layer security, invoked based on-site policy, supports the highest level of filtering. Individual commands within applications, as well as access to specific machines and services, can be permitted or denied. Application layer mechanisms are discussed in the opening part of the VPN Section 5.3. One of the major shortcomings of application layer mechanisms is that they rely on platforms with minimal trust mechanisms and that connections must be established at a lower level in the protocol stack (network and transport layer) before the application mechanisms are applied. This leaves the machine vulnerable to network attacks that are unaffected by higher-layer security mechanisms. The other drawback of application layer security is the number of applications that need to be covered. As application protocols evolve, security is usually a secondary consideration. The number of application software packages offered in the commercial market (for example, e-mail packages) makes it difficult to add security services to every package as a retrofit. Efforts to standardize the interface to security services will help this problem, but are ineffective if the vendor is simply not interested in implementing security services in the product.

## 6.2.6 Cases

This version of the Framework does not address remote access of Top Secret or higher sensitivity level information. By definition, the disclosure of this information can cause exceptionally grave damage to national security. Remote access to Top Secret information presents extreme risk and should be handled on a case-by-case basis.

This section considers remote access to information at the SBU level and the remote access to classified information up to the Secret level as separate cases. Secure remote access to Top Secret information may be addressed in future versions of this document.

As depicted in Figure 6.2-5, the two different access paths combined with the two sensitivity levels produce four generic cases: Secret dial-in access, Secret ISP access, SBU dial-in access, and SBU ISP access. For each case, the underlying network options include PSTN, ISDN, and other digital and wireless services.

The specific requirement cases include the following.

- Remote access to Secret enclave via direct connection through PSTN, ISDN, Wireless connections, and other digital connections.
- Remote access to Secret enclave via ISP connection through PSTN, ISDN, Wireless connections, and other digital connections.
- Remote access to Unclassified enclave via direct connection through PSTN, ISDN, Wireless connections, and other digital connections.
- Remote access to Unclassified enclave via ISP connection through PSTN, ISDN, Wireless connections, and other digital connections.

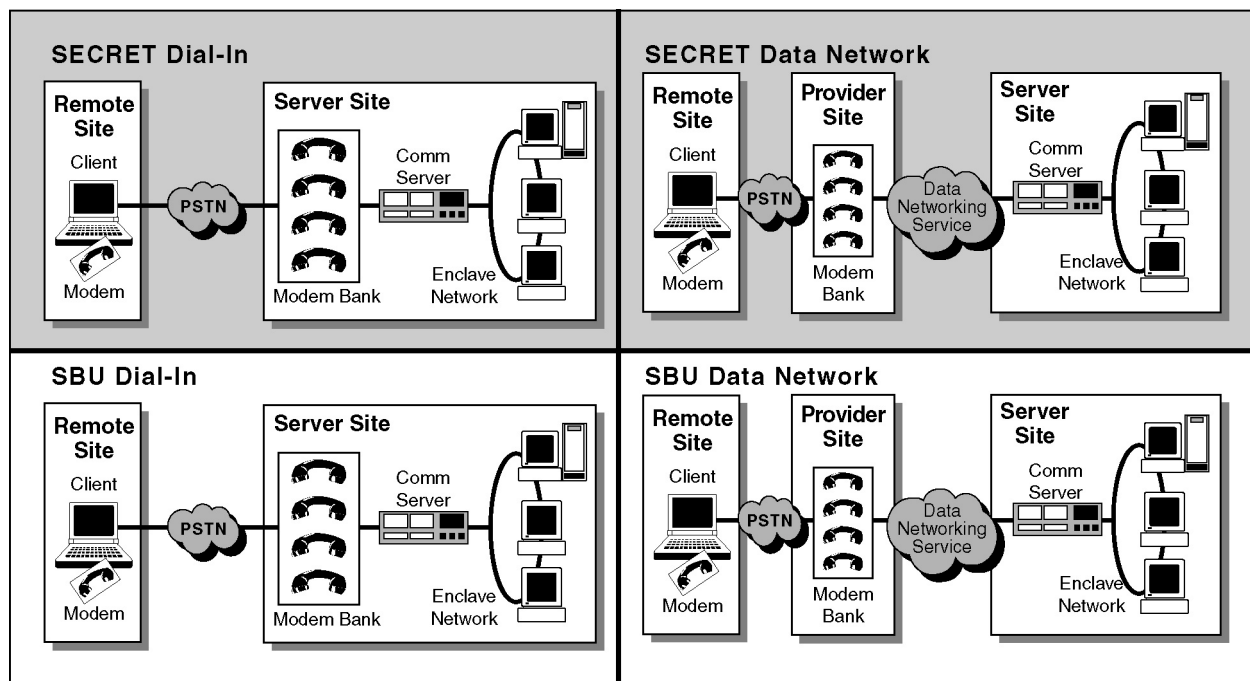


Figure 6.2-5 Remote Access Cases

## 6.2.7 Framework Guidance

The following guidance is based on the premise that the home site has properly followed an information systems security engineering process. This process will identify the organization's assets and vulnerabilities and provide a total system solution that mitigates the risk to the level decided by the organization. The discussion here is at a generic level. The level of risk acceptance and the availability of products and services will determine a site's remote access security solution.

### 6.2.7.1 Case 1: Remote Access to Secret Enclave via Direct Connection over PSTN

Guidance for this case is summarized in Table 6.2-1. Each of these tables is followed by a discussion of the rationale behind the recommendations.



## Media Encryption

A media encryptor is recommended to protect the information stored in the remote user computer. The rationale for this is that media encryption provides confidentiality for data on the user's hard drive. It also performs a workstation integrity function by protecting the integrity of the computer's configuration (e.g., by verifying the BIOS and making sure that the user is notified of any modifications to applications and hardware configuration files).

**Table 6.2-1a Summary Guidance for Remote Access  
Direct Dial-up Access to Secret Enclave**

Primary Solution Components	Guidance Categories	Desired Solution	Best Commercially Available Solution	Gap Between Needed & Available Solution
<b>Media Encryptor</b>	Role of this component	To protect the confidentiality and integrity of all data stored on the hard disk in the event that the user's laptop is lost, stolen, or tampered with. To keep the laptop unclassified when not in use		
	Security functions	Dynamically encrypt all data (but system boot files) stored on the hard disk  Protect the private key used to encrypt the data by storing it on a token that is physically removed when not in use  Require user PIN to unlock the token	Hardware token-based, software media encryption for Windows platforms	WIN95 and WIN NT versions
	Cryptographic strength (if applicable)	Crypto algorithm and key length should be of robustness level 2.	Type II algorithm (SKIPJACK) w/ 80 bit key	TBD
	Common Criteria Assurance Level	EAL 4	N/A	Three assurance levels
	SMI/PKI/KMI services	Generation of file encryption keys  Data recovery in event of lost token or user PIN		
	SMI assurance	KMI level 2	TBD	TBD
	Interoperability requirements	No requirement	No commercial standards exist. Current solutions are not compatible with each other	Interoperability

The remote computer needs certain system files in order to boot, so these files should remain unencrypted on the storage media. However, the proper functioning of the media encryptor depends on the integrity of the boot process, so the integrity of these unencrypted system files must be verified. The media encryptor should also verify the integrity of the computer's BIOS

configuration. All other space on the storage media should be encrypted. The media encryptor should verify the system's integrity upon boot-up and notify the operator if integrity checks fail. The media encryptor should use algorithms approved for the protection of Secret information. To help mitigate concerns about weak or compromised keys, the media encryptor should be capable of accepting keys from an outside source (e.g., FORTEZZA® card and its associated security management infrastructure). The implications of having a split-key are discussed in Chapter 8, Supporting Infrastructures of this Framework. The media encryptor should support two roles: user and system administrator. Only the system administrator should have the ability to change the configuration of the remote computer and the media encryptor. Depending on the user's environment and the organization's security policy, the media encryptor could also be used to preclude the booting of the remote computer via an unencrypted floppy disk. If the remote user wants to access unclassified systems, it is recommended that the user utilize a separate hard drive for this purpose, since the costs of implementing and maintaining a trusted operating system (to maintain data separation and integrity) would typically be prohibitive.

## Remote Workstation Integrity

Recommendations concerning remote workstation integrity are contained in the Firewalls section, Section 6.1.6 and are summarized here. Enclave boundary and protection components should be chosen in accordance with the site's security policy. The user's home enclave should choose a network boundary protection mechanism (e.g., guards, firewalls) paying close attention to the tradeoffs among security, performance, and cost. An intrusion detection system may be implemented, but technology in this area is not mature. A virus scanning policy should be implemented, with scans occurring periodically or after certain events. Network vulnerability scanners should be run periodically, and identified deficiencies should be addressed.

**Table 6.2-1b Summary Guidance for Remote Access  
Direct Dial-up Access to Secret Enclave**

Primary Solution Components	Guidance Categories	Desired Solution	Best Commercially Available Solution	Gap Between Needed & Available Solution
<b>Workstation Integrity</b>	Role of this component	Protect the remote user's workstation against unauthorized modification		
	Security Functions	Digital signature and integrity hash function	Digital Signature Standard and Secure Hash Algorithm	
	Cryptographic strength (if applicable)			
	Common Criteria Assurance Level	EAL4	N/A	Three Assurance Levels
	SMI/PKI/KMI Services			
	SMI assurance			
	Interoperability requirements			

Remote user and enclave software should be kept up-to-date, since many discovered vulnerabilities are patched in later versions. In addition, software should be protected from tampering by cryptographic checksums applied at the manufacturer and should be checked when the software is installed (on the user's workstation or the enclave components). New versions of software could also inject new vulnerabilities into the system and thus should be tested before operational use.

Other mechanisms used to protect the integrity of the remote user's workstation include trusted operating systems, hardware tokens, user password authentication, and so on. At least in the case of a Secret enclave, the remote user should be afforded the same protection mechanisms that are provided to the user's workstation located in the user's home enclave. In addition, the user's environment will dictate extra security services, as required by the organization's security policy. For instance, special policy and procedures are typically required in higher threat environments in which physical security is not at the same level as provided at the home enclave. Additional security mechanisms should give the user the tools to mitigate the loss of workstation integrity.

**Table 6.2-1c Summary Guidance for Remote Access  
Direct Dial-Up Access to Secret Enclave**

Primary Solution Components	Guidance Categories	Desired Solution	Best Commercially Available Solution	Gap Between Needed & Available Solution
<b>Secure Modem</b>	Role of this component	Authenticate and encrypt the connection between the remote user and the home enclave		
	Security Functions	Mutual authentication Continuous authentication Full period encryption at the secure modem layer In-line encryption Hardware device Removable hardware token to store and protect private keys User PIN to unlock token	Encrypting modem supporting KEA and SKIPJACK	
	Cryptographic strength (if applicable)	Secret	Secret w/ NAG-68 Interim Policy	Secret
	Common Criteria Assurance Level	EAL3	N/A	Three Assurance Levels
	SMI/PKI/KMI Services			
	SMI assurance	KMI level 2	TBD	TBD
	Interoperability requirements	Support for AT command set and communications protocol standards Software compression	56Kbps.X.90	Interoperability

## Enclave Boundary and Connection Protection

A link-encrypting device should be used to protect the communications link between the remote user and its home classified enclave. To be used in a classified environment, the device must provide strong authentication and confidentiality services. Modems should meet the applicable commercial standards, such as V.nn and MNPnn. The modem should provide an AT commands interface. To authenticate the remote user to the modem, the modem should require the entry of a Personal Identification Number (PIN) to enable the encrypted data mode. The modem must pass I&A information to the boundary protection mechanism for system access (See Section 6.2.5, Technology Assessment). GUI software should be provided to allow the entry of the PIN and it should display authenticated identities and security modes of operation. The modem may have a plaintext mode of operation (other than that required by the initial “handshaking” done before a secure session is established). Use of this mode should require overt action on the part of the user so that this mode is not selected by accident or by default. Explicit requirements for secure modems will be provided in later releases of the Framework.

In addition to the encrypting modem, a boundary protection device should identify and authenticate the dial-in user at the point of presence of the classified network to the local PSTN. This is discussed in more detail in the next section.

**Table 6.2-1d Summary Guidance for Remote Access  
Direct Dial-up Access to Secret Enclave**

Primary Solution Components	Guidance Categories	Desired Solution	Best Commercially Available Solution	Gap Between Needed & Available Solution
<b>Enclave Boundary Protection</b>		Mutual and continuous authentication Full period encryption at the secure modem layer In-line encryption Hardware device User PIN to unlock token	Secure communications server supporting encrypting modem	
<b>Solution Residual Risks</b>		<b>None</b>	<b>Acceptable</b>	<b>Difference</b>

## Authentication Mechanism

An additional authentication mechanism should be implemented that will provide strong authentication directly to the boundary protection mechanism to implement a “that which is not explicitly permitted is denied” policy. For example, many remote users only need e-mail while they are traveling; in addition, some may need access to a particular file server. Providing the minimum access needed to do the job not only mitigates the effects of any successful attack by an outsider, but also makes insider attacks more difficult. Guards and firewalls provide this functionality.

Authentication to the user's workstation is recommended. A password, hardware/software token, or biometric device should be used, depending upon the level of assurance required. See the Firewall, Section 6.1 for more information on this issue.

## Technology Gaps

The only Government-Off-The-Shelf (GOTS) solution supporting the remote access user is the AT&T Secure Telephone Unit (STU)-III 1910 Secure Data Device (SDD). The SDD runs at data transfer rates much lower than those of modems available in today's commercial market. A cumbersome device, the 1910 is actually heavier and larger than the laptop it supports. There is a consensus in the user population that there is no technology available today. No technology currently provides a high enough level of assurance to pass classified data over the PSTN to and from a classified enclave at the same level of performance that is available in non-encrypting Commercial-Off-The-Shelf (COTS) modems. This gap is certainly noticeable when comparing capabilities with the 56 Kbps modems on the market today.

In general, there is a technology gap in high-assurance security solutions applicable to remote access in the COTS environment. In particular, little commercial work is being done on media encryptors, although several file encryption products are available. File encryptors are not widely available for non-Windows operating systems. A few commercial encrypting modems are available, but high-assurance encrypting modems are not commercially available. In addition, secure remote access servers and communication servers are not widely available. Support for Top Secret remote access will require additional features that are not available in today's commercial marketplace, at least at an acceptable risk level. Workstation integrity and configuration guidance are also issues. Future versions of this Framework will address these gaps in more detail.

### 6.2.7.2 Case 2: Remote Access to Secret Enclave via ISP Connection

This section will be provided in a future release of the Framework.

### 6.2.7.3 Case 3: Remote Access to Unclassified Enclave via Direct Connection

The recommended solution for this case involves implementing a RADIUS server within the enclave and configuring each remote workstation with a RADIUS client. When a remote workstation requests access to the network, RADIUS based authentication is used.

- a) *Media Encryption*—In this scenario, all information is Unclassified. Therefore media encryption is not necessary for information stored on the remote workstation. File encryption may be desired for protection of SBU information.
- b) *Workstation Integrity*—An Unclassified remote access workstation will also likely have access to the Internet. There may be a requirement for the remote workstation to download files from the Internet or to exchange files with the Unclassified enclave.

Downloading files from the Internet poses a risk to the workstation's integrity. The workstation should have a robust and updated virus scanning capability. Additionally, the workstation connecting to the enclave poses a risk to the integrity of the enclave if precautions are not taken to check for viruses on the workstation. Again, to protect the integrity of the workstation and the enclave, virus scanning should be resident on the remote workstation.

- c) *Enclave and Connection Protection*—The enclave is vulnerable to unintentional virus insertion through the remote workstation. Although RADIUS-based authentication of remote workstations prevents unauthorized remote workstations from gaining access to the enclave's network, there is still a risk of valid workstations being lost or compromised.

All workstations should be equipped with a robust user-to-workstation authentication mechanism. Although in the case of workstation theft or compromise, this mechanism alone may not provide adequate assurance that the workstation cannot be used to access the enclave. A way of mitigating the risk of such access is by implementing an incident report procedure for reporting lost or compromised remote workstations and by installing and maintaining an intrusion detection system. If a lost or compromised workstation is reported in a timely manner, the RADIUS server can be configured to deny access from that compromised workstation. If the compromised workstation establishes a connection to the network before the compromise is reported and mitigated, an intrusion detection system will identify anomalous behavior and alert administrators to the possibility of a compromised workstation.

Although the user information in this scenario is Unclassified, there still may be a requirement to provide confidentiality for the connection. A VPN solution can be established across the remote connection. A layer 2 mechanism, such as L2TP, or a layer 3 mechanism such as IPSec may be implemented to provide confidentiality. These technologies are discussed in further detail in Section 5.3.

- d) *Authentication Mechanism*—Authentication between the remote workstation and the home enclave is achieved by using the RADIUS protocol. The RADIUS protocol relies on a shared secret between the RADIUS client and the RADIUS server. MD5 is used to hash the shared secret, the user password, and other fields in the RADIUS message. The strength of the authentication is based on protecting the shared secret.

Authentication to the user's workstation is also recommended. A password, hardware/software token, or biometric device should be used, depending on the level of assurance required.

#### 6.2.7.4 Case 4: Remote Access to Unclassified Enclave via ISP Connection

The recommended solution for this scenario involves implementing an IPSec-compliant firewall or other boundary protection device. Remote workstations must be configured with an IPSec-compliant network card, software, or other component. This case also involves implementing a RADIUS server within the enclave and configuring each remote workstation with a

RADIUS client. In this scenario, the remote workstation must usually still use the PSTN to establish a connection to the ISP. The ISP then interfaces with the Internet, which interfaces with the enclave. The remote workstation establishes an IPSec-secured connection over the PSTN that terminates at the enclave ISP-compliant firewall or boundary protection device.

*Media Encryption*—In this scenario, all user information is Unclassified. Therefore, media encryption for information stored on the remote client is not necessary. File encryption may be desired for protection of SBU information.

- a) *Workstation Integrity*—An Unclassified remote workstation will also likely have access to the Internet. There may be a requirement for the remote workstation to download files from the Internet or to exchange files with the Unclassified home enclave. Downloading files from the Internet poses a risk to the workstation's integrity. The workstation should have a robust and updated virus scanning capability. Additionally, the workstation dialing connecting to the enclave poses a risk to the integrity of the enclave if precautions are not taken to check for viruses. Therefore, to protect the integrity of the workstation and the enclave, virus scanning should be resident on the remote workstation.
- b) *Enclave and Connection Protection*—The enclave is vulnerable to unintentional virus insertion through the remote workstation. Although RADIUS-based authentication of remote workstations prevents unauthorized remote workstations from gaining access to the enclave's network, there is still a risk of valid workstations being lost or compromised.

All workstations should be equipped with a robust user to workstation authentication mechanism. Although in the case of workstation theft or compromise, this mechanism alone may not provide adequate assurance that the workstation will not be used to access the enclave. A way of mitigating the risk of such access is by implementing an incident report procedure for reporting lost or compromised remote workstations and by installing and maintaining an intrusion detection system. If a lost or compromised workstation is reported in a timely manner, the RADIUS server can be configured to deny access from that compromised workstation. If the compromised workstation succeeds in establishing a connection to the network before the compromise is reported and mitigated, an intrusion detection system will identify anomalous behavior and alert administrators to the possibility of a compromised workstation.

Although the user information in this scenario is Unclassified, there still may be a requirement for confidentiality. If confidentiality is required, the IPSec client on the remote workstation can use the ESP feature of IPSec to encrypt the IP payload.

- c) *Authentication Mechanism*—Authentication between the remote workstation and the home enclave is achieved by using the authentication header of IPSec. The IPSec authentication header relies on a shared secret using either a symmetric encryption algorithm (i.e., Data Encryption Standard (DES)), or a one way hashing algorithm (i.e., MD5, HA).

Authentication to the user's workstation is also recommended. A password, hardware/software token, or biometric device should be used, depending on the level of assurance required.

This page intentionally left blank



## 6.3 Guards

Guards enable users to exchange data between private and public networks, which is normally prohibited due to information confidentiality. The combination of hardware and/or software guards are used to allow secure Local Area Network (LAN) connectivity between enclave boundaries operating at different security classification levels (i.e. one private and the other public). Guard technology can bridge across security boundaries by providing some of the interconnectivity required between systems operating at different security levels. Several types of guards exist. These protection approaches employ various processing, filtering, and data-blocking techniques in an attempt to provide data sanitization (e.g., downgrade) or separation between networks. Some approaches involve human review of the data flow and support data flow in one or both directions. Information flowing from public to private networks is considered an upgrade. This type of transfer may not require a review cycle, but should always require a verification of the integrity of the information originating from the public source system and network. This section will discuss guards, the environment and mannerism in which they are most suited for implementation, how they can be used to counteract attacks made on the enclave as well as the different variety of guards and their functions.

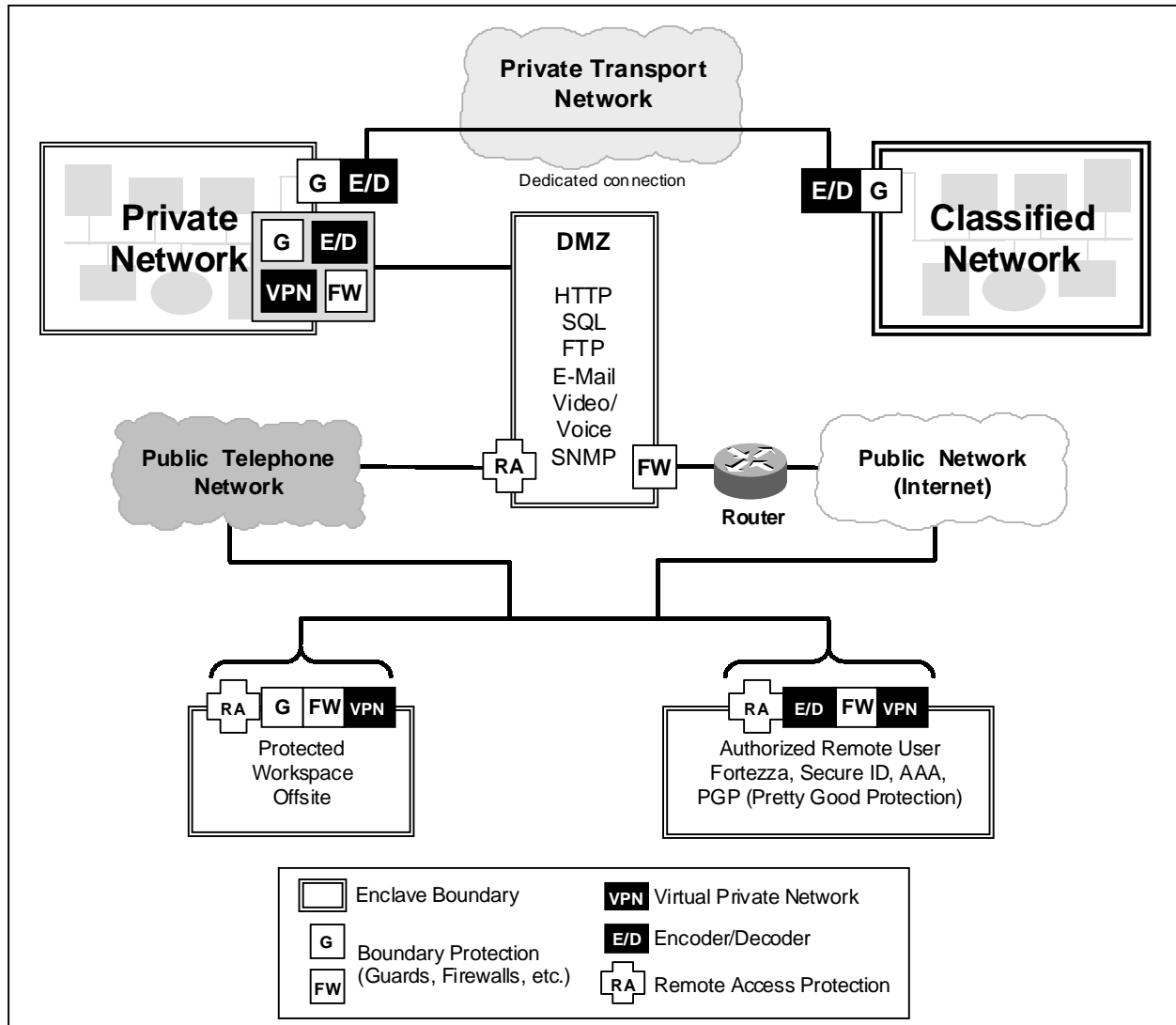
A guard is a device used to defend the network boundary by employing the following functions and properties.

- Typically subjected to high degree of assurance in its development.
- Supports fewer services.
- Services are at the application level only.
- May support application data filtering (review).
- May support sanitization of data.
- Typically used to connect networks with differing levels of trust (provides regrading of data).

### 6.3.1 Target Environment

The guard is designed to provide a secure information path for sharing data between multiple system networks operating at different security levels. The overall system that employs a guard is illustrated in Figure 6.3-1 and is comprised of a server, workstations, malicious code detection, a firewall, and/or filtering routers all configured to allow transfer of information among communities of users operating at different security levels. The server and workstation components may implement a hardware- or software-based authentication scheme to authenticate to the guard. The firewall component is usually Commercial Off-The-Shelf (COTS) hardware and/or software that filters the network traffic and is configured to forward only authorized packets. A commercial filtering router may also be used to perform this function. The firewall's primary function is to provide barriers against successful penetration of the low side LAN by unauthorized external users. The firewall hides the networks behind it and supplements the guard. The firewall restricts access to all traffic other than the traffic being scrutinized by the guard. Virtual Private Networks (VPNs) can also be employed using either a

firewall or other encryption device. To ensure the security of the overall system, all users, managers, and system administrators must exercise the security policies and practices of the organization. Some considerations include: valid personnel approval for access to all information stored and/or processed on the system; formal access approval process for, and signed non-disclosure agreements for all information stored or processed on the system; valid need-to-know process for some of the information stored or processed by the system. Communication links, data communications and data networks of the system must protect the network determined by the sensitivity level of data on that particular network.



**Figure 6.3-1 Guard Environment**

The guard can be configured to function in different directions.

- The private to public bi-directional mode facilitates data to move from private to public after the review process for releasability to the lower network classification. Data moving from low to high, need not undergo the review process for releasability, but processing, filtering, and blocking should occur to identify viruses and other malicious code

transfers. Users on the private network would be allowed to push public data to users on the public network, and in turn, users on the public network could push public data to users on the private network. Users on the private network would also be allowed to view and pull data that exists on the public network.

- The private to public unidirectional mode allows data to move from private to public after the review process for releasability to the lower network classification. No transfer is permitted from the lower network to the private network. Users on the private network would send data to be downgraded to the public level, which would then be pushed to a server on the public network for subsequent pull by users on the public network.
- The peer to peer mode allows communications between networks bridged by the guard at the same security level (e.g., private and private-releasable). That is, all the screening the guard normally does on private to public transfers in the private/public configuration is performed in both directions. Standard operating procedures must be implemented so that appropriately cleared personnel from each side can administer the guard screening criteria databases. This configuration allows users on the private network to downgrade data to the private-releasable level, and push that data to a server on the private-releasable network for subsequent pull by users on the private-releasable network.

## 6.3.2 Requirements

This section addresses the functional requirements of the communication, releasability, and network access capabilities.

### 6.3.2.1 Communication Requirements

Requirements for communication include the following.

- The guard shall allow users on the private networks to communicate with only specified hosts on the public networks.
- The guard shall prohibit workstations to be used as a pass-through or gateway device from either the private or public sides for any communications, including mail.
- The guard shall send public data to one of the public networks or private networks using the appropriate router.
- Routers shall be configured to restrict the types of network services that may pass through them as well as the sources and destinations of service requests.
- The guard shall transfer the appropriate data from the private network to the public network.
- The guard shall allow protocols to pass through it.
- The guard shall allow only authorized users to send and/or receive a message by performing access control on both the source and destination addresses of the message.

## 6.3.2.2 Releasability Requirements

Current requirements for releasability include:

- The guard shall allow only a properly labeled message to pass from the private level to the public level.
- The guard shall support a policy that allows only attachments that have been reviewed for security level at the user's workstation to pass from the private-to-public side.
- The guard shall allow only selected application attachments to pass through it—this capability will be configurable to support a variety of application packages.
- The guard shall perform word and/or phrase search.
- The guard shall support rule-based sanitization (i.e., message content modification) of messages from high levels through low levels.
- The guard shall ensure that only allowed data is distributed.
- The guard shall validate proper message construction, including configurable verification of message content.
- The guard shall remove classification labels, which were inserted into the e-mail body and attachments prior to delivery to the other side.

## 6.3.2.3 Access Requirements

Current access requirements for file transfers include:

- The guard shall run on a trusted platform.
- The guard shall prevent message flow directly between the private side Wide Area Network (WAN) and the guard in either direction.
- The guard shall support a programmable set of security ID labels per flow.
- The guard shall ensure that the security level of a message subsumes (is equal to or greater than) the security level of its attachment(s).
- The guard shall protect against unauthorized disclosure of private side information.
- The guard shall provide safeguards to protect the private side from attacks (including penetration, malicious code, and denial of service) from the public side.
- The guard shall support user authentication and encryption capabilities.
- The guard shall perform audit all security-related functions.
- The guard shall provide an access control mechanism to limit access to the controls and provide separate roles for the security administration, system operator, and mail administration functions. Thus, a supporter authorized to function in one area will be prevented from performing functions in another, unless specifically given permission to do so.
- The guard shall prevent disclosure or release data to unauthorized consumers.
- The guard shall provide a secure bridge for passing messages between networks of differing levels of security.

- The guard shall strip off the digital signature as the message passes through the guard.
- The guard shall restrict source routing. Source routing is a form of addressing which can alter the routing of a message from its normal route.
- The guard shall journal/log all passed and/or failed messages.

## 6.3.3 Potential Attacks

The focus within this category is on attacks into an enclave by malicious e-mail, file, or message transfers. Guards can be implemented to provide a high level of assurance for networks by preventing certain types of malicious messages from entering the enclave. The types of attacks are categorized into three sections: Section 6.3.3.1 Active Attacks, Section 6.3.3.2 Distribution Attacks, and Section 6.3.3.3 Insider Attacks. For more information related to attacks, please refer to Chapter 4.2-Adversaries, Threats (Motivations/Capabilities), and Attacks.

### 6.3.3.1 Active Attacks

Active attacks include those attacks that attempt to breach security features or exploit data in transit whether it be e-mail, file or message transfers. Some firewall technologies and e-mail systems that perform content filtering will help establish a level of trust for messages that are signed but not encrypted. Messages may be signed and/or encrypted at the user level and/or the organizational level. However, a digital signature on a message does not increase the safety level for the contents of the message. Active attacks can include, but are not limited to, the insertion of malicious code or the theft of data. Examples of active attacks in regards to the transmission of messages and files are listed below. For further description of network-based attacks, please refer to Section 4.2.1.4.2, Network Based Vulnerabilities and Active Attacks.

- Modification of Data in Transit – Modifications are not necessarily always malicious or intentional. A modification could be the conversion of spaces to tabs or vice versa within an e-mail or real-time message. A network-based modification could also be the occurrence of a complete violation of standards. Internet e-mail standards necessary for the secure transmission of messages from one domain to another include PGP-Pretty Good Privacy, MIME – Multipurpose Internet Mail Extensions, and S/MIME – Secure Multipurpose Internet Mail Extensions. While instant/real-time messaging do not yet have interoperable standards established, protocols must be established to ensure that the messages have not been intercepted and corrupted.
- Insertion of Data – Reinsertion of previous messages.
- Inserting and Exploiting Malicious Code (Trojan Horse, Trap Door, Virus, Worm).
- Defeating login mechanisms into e-mail accounts, messaging accounts or into file storage servers.
- Session Hijacking – In the case of e-mail, file or real-time message transfers unauthorized access could be gained into a communications channel with malicious intent.
- Denial of service.
- Establishment of unauthorized network connections.

- Masquerading as an Authorized User – An attacker would use the identification of a trusted entity to gain unauthorized access to information either through e-mail, real-time messaging or by requesting file transfers.
- Manipulation of data on the private side.
- Decrypting weakly encrypted traffic.
- Misrepresentation or information “faking” through Internet Relay attacks – Third-party mail relay occurs when a mail server processes and delivers e-mail from an external client. In this manner, mail appears to originate from that mail server's site and not the original site. Spam e-mail is generally distributed this way, at the mail owner's expense. Intruders can spam e-mails with embarrassing content or by flooding a site with e-mails. Damage caused by spamming includes the loss of reputation of the system that has been identified with the attack e-mail as well as the loss of connectivity to large parts of the Internet that have blocked sites from spamming. E-mail servers will become clogged, mail can be lost or delivered late, and then there are the cleanup costs to remove spammed mail without destroying legitimate mail. For more information, please refer to <http://www.ciac.org/ciac/ToolsTestEmail.html>. [1]
- Monitoring Plain Text Messages – Plain text messages are those messages that are not encrypted, and therefore not secure in any manner. Once intercepted, plain text messages can be easily read.

### 6.3.3.2 Distribution Attacks

Distribution attacks can occur anytime during the transfer of a Guard's software and/or hardware. The software or the hardware could be modified during development or prior to production. The software is also susceptible to malicious modification during production or distribution. Section 6.3.4.2 will discuss methods in which these attacks could be prevented. For additional information, please refer to Section 4.2.1.4.4, Hardware/Software Distribution Vulnerabilities and Attacks. Also, refer to Table 4-3, Examples of Specific Modification Attacks.

### 6.3.3.3 Insider Attacks

While an enclave must be protected from outside intruders, it must also be protected from attacks from inside the enclave. Interception or attacks to messages can occur during transit from the insider level. The originators' and recipients' mail system administrators have the ability to look at e-mail messages and files that are being sent. E-mail messages that bounce back usually have a copy sent to the e-mail system administrator to help determine the reason behind the bouncing, therefore the administration has messages that have bounced brought to his/her attention with full viewing privileges to the message that is attempting to be sent. An insider attack occurs when someone located within the boundaries of the enclave intercepts or modifies data or security mechanisms without authorization.

Unauthorized access could also be gained into the overhead portion of a covert channel. The use of a covert channel is a vulnerable point of attack due to the transport overhead not being completely defined and therefore susceptible to exploitation. The physical theft of data is another threat within the enclave. For further detail, please refer to Section 4.2.1.4.3 – Insider Vulnerabilities and Attacks.

## 6.3.4 Potential Countermeasures

For all efforts aimed at attacking an enclave through the unauthorized access or modification to e-mail messages, real-time message transfers, or file transfers there must be measures in place to prevent these attacks from penetrating the boundaries of an enclave. In the case of attacks that originate from inside the enclave, precautionary measures also need to be taken in areas that are vulnerable to attacks including the physical theft and unauthorized access to data. The following subsections address measures that can be taken to counteract attacks against an enclave and information transfers among enclaves. These countermeasures are placed into three categories: Boundary Protection via Guards, Distribution Attack Countermeasures, and Insider Attack Countermeasures.

### 6.3.4.1 Boundary Protection via Guards

Guards can be implemented to protect the enclave and the messages passing within and through the enclave boundaries. Guards enable users to exchange information between either networks of the same or differing classification levels. Traffic analysis is a means by which traffic can be monitored. Traffic analysis can be conducted to help identify traffic patterns, i.e., origination and destination endpoints for traffic, and thus aid in the discovery of the endpoints of unauthorized network connections. Enclave boundaries need protection from the establishment of unauthorized network connections. The responsibility lies with the management and administration of the local network to prohibit unauthorized connections between networks of different classification levels and to enforce this policy through non-technical means.

The following bulleted items list the type of attack and the countermeasure that can be used to prevent that attack from occurring.

- **Modification of Data in Transit**—The countermeasure to this attack is to use digital signatures or keyed hash integrity checks to detect unauthorized modification to the data in transit. E-mail, real-time message transfers as well as file transfers are all susceptible to interception and modification while in transit.
- **Insertion of Data**—There are many countermeasures to the malicious insertion of data. They include the use of timestamps and sequence numbers, along with cryptographic binding of data to a user identity, in order to prevent the replay of previously transmitted legitimate data. Data separation or partitioning techniques, such as those used by guards and firewalls, deny or restrict direct access and the ability to insert data during transit.
- **Insertion and Exploiting Malicious Code (Trojan Horse, Trap Door, Virus, Worm)**—Implement a guard and employ strong authentication in order to filter and block incoming messages that are not from authenticated parties. To help ensure that mail is not modified during transit or forged, technologies and products such as PGP and S/MIME can be used to encrypt and sign messages on a regular basis. Real-time messaging protocols are necessary to also ensure authentication amongst parties.
- **Defeating Login Mechanisms**—The most appropriate countermeasure for this is the cryptographic authentication of session establishment requests. This pertains to logging into an e-mail account or for obtaining access to a file server or messaging channel.

- Session Hijacking—The countermeasure for this is continuous authentication through digital signatures affixed to packets, or at the application layer, or both.
- Denial of Service—Countermeasures that can be taken against these attacks include having a guard to filter out bad source Internet Protocol (IP) addresses, filter Internet Control Message Protocol (ICMP) echo responses or limit echo traffic, guard against all incoming User Datagram Protocol (UDP) service requests. A non-technical countermeasure would be to subscribe to the Certification and Accreditation (CERT) mailing list ([www.cert.org](http://www.cert.org)) in order to receive notifications every time a new Internet weakness emerges. [2]
- Establishment of Unauthorized Network Connections—A non-technical countermeasure lies with the management and administration of the local network to prohibit and enforce the policy against unauthorized connections between networks of different security levels. Commercial tools are also available for system administration personnel to use for detecting unauthorized connections. Unauthorized connections would allow for otherwise prohibited access to e-mail and data files as well as real-time message interception.
- Masquerading as an Authorized User—The appropriate countermeasure is to use cryptographic authentication in conjunction with timestamps or sequence numbers to prevent any recording and/or replay of authentication data whether it be e-mail, real-time messaging or file transfers. Another countermeasure to prevent stealing an authentic session is to cryptographically bind authentication data to the entire session/transaction.
- Manipulation of Data on the private Side—The appropriate countermeasure is to permit only authorized users to access the data, through file transfers, on the private side using cryptographic authentication and data separation techniques.
- Decrypting Weekly Encrypted Traffic—To ensure that unauthorized persons cannot access e-mail messages, real-time messages, or files in transit, adequate encryption algorithms and sound key management processes must be observed
- Misrepresentation or information “faking” through Internet Relay attacks—The countermeasure for these spamming attacks would involve the use of a guard to filter the messages and therefore block malicious messages, whether they are e-mail messages or real-time messages, from entering the enclave.
- Monitoring Plaintext Messages—The monitoring of messages can be counteracted by denying access to the data by unauthorized users. Access denial is possible by encrypting the data or by using other data separation techniques that will restrict those who are unauthorized from obtaining access to the data contained within a file.

### 6.3.4.2 Distribution Attack Countermeasures

During the development, manufacturing, and distribution stages, both technical and non-technical measures must be taken to avoid the malicious modification of guard software and hardware. The following lists the stage at which an attack could occur and the countermeasure to prevent such an attack.



- **Modification of Software or Hardware during Development, Prior to Production**—Strong development processes/criteria are essential during this phase as a countermeasure for threats. Continuous risk management through processes, methods, and tools is also necessary. The following web site link contains a collection of software engineering processes, methods, tools, and improvement references, <http://www.sei.cmu.edu/managing/managing.html>. [3] Subsequent third party testing and evaluation of software should also be conducted to ensure that the software/hardware has not been modified. High assurance methods and criteria should be followed, such as the Trusted Product Evaluation Program (TPEP) and Common Criteria, please refer to <http://www.radium.ncsc.mil/tpep/tpep.html> for program details. [4]
- **Malicious Software Modification During Production and/or Distribution**—The countermeasures for threats during this phase include high assurance configuration control, cryptographic signatures over tested software products, use of tamper detection technologies during packaging, use of authorized couriers and approved carriers, and use of blind-buy techniques.

### 6.3.4.3 Insider Attack Countermeasures

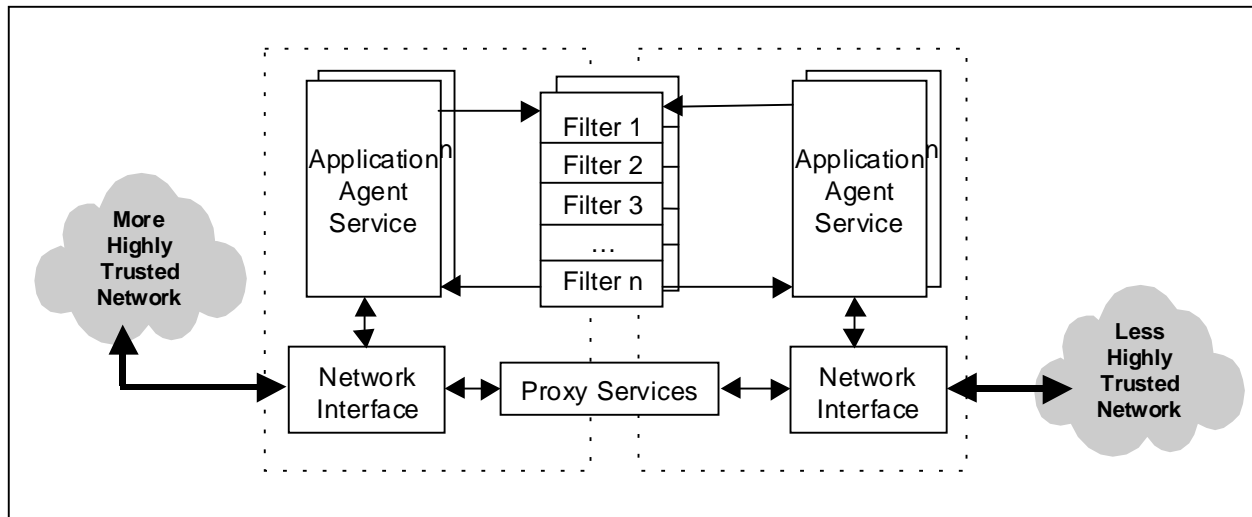
Technical and non-technical countermeasures must both be taken to prevent against attacks originating within the boundaries of an enclave. The following are the types of insider attacks that can occur and the countermeasure that must be taken to prevent the attack.

- **Modification of Data or Modification of Security Mechanisms by Insiders**—The primary technical countermeasure is to implement auditing procedures of all actions taken by users that could pose a threat to security. Audit logs will need to be generated and timely, diligent reviews and analysis must be conducted. Non-technical countermeasures include personnel security and physical procedures.
- **Physical Theft of Data**—Appropriate non-technical countermeasures include personnel security and physical security procedures, which inhibit actual removal of data, either in printed form or on storage media.
- **Covert Channels**—The countermeasure against a covert channel between networks of different classification levels is a trusted guard function that examines network header fields and network messages for possible unauthorized information.

### 6.3.5 Guard Technology Assessment

Guards are usually used to enable connectivity that is normally prohibited because of information confidentiality. Where a firewall is usually used to restrict or scrutinize information flow on an already existing link to LAN or WAN circuits, guards allow the transfer of information between segments. A combination of hardware and software components designed to allow secure Local Area Network (LAN) connectivity between segments operating at different security classification levels (one private and the other public). The vast majority of guard implementations use a dual network approach, which physically separates the private and public sides from each other. Please refer to Figure 6.3-2 below. As the figure also depicts, guards are application-specific and therefore all information will enter and exit by first passing through

the Application Layer, Layer 7, of the OSI Model. In addition, most guard processors are high assurance platforms that host some form of trusted operating system and trusted networking software.



**Figure 6.3-2 Dual Network Approach**

Guards can be a fully automated (without any human intervention) Multi-Level Security guard system that permits one-way or bi-directional transfers of data between multiple LAN systems operating at different security or releasability levels. Guards can concurrently review and sanitize multiple binary and ASCII files, and virtually any complicated data format. Almost any data type that can be “packaged” into a file can be transferred through certain guards, including SQL, HyperText Transfer Protocol (HTTP), UDP, SMTP/e-mail attachments, and others. The guard controls the automated information flow between multiple LAN systems according to security rule filters. When implemented in conjunction with a firewall a higher degree of security for protecting the enclave is achieved.

This section is further broken down to discuss guard technological areas that can be used to protect the enclave:

- 1) Authenticated Parties Technologies.
- 2) Confidentiality and Integrity.
- 3) Data Processing, Filtering, and Blocking Technologies.

This categorization allows for a high level assessment of system assurance so that a determination can be made as to the level of security robustness a network will require. These three categories of potential protection approaches are explained in more detail in the following subsections.

### 6.3.5.1 Authenticated Parties Technologies

Approaches for protecting the enclave that are included within this category are those which mandate the use of cryptographic authentication mechanisms prior to allowing access. Authentication allows two parties that intend to exchange data to identify themselves to one

another and positively authenticate their identities. Hence, they become mutual trusting parties. The data that flows between these trusting parties is at the lower security level. Authenticated access is widely available and is supported by a large number of standards and protocols. Authentication protects the enclaves of private users that are separated from Public network users through an enclave boundary protection device such as a guard and/or Firewall. In such a topology, Public network users might utilize digital signature technology to authenticate themselves to private network users. In addition, the guard might incorporate Access Control List (ACL) mechanisms to make access decisions governing the set of users that are authorized to release information from the private network. Access control lists can also be used to restrict the set of Public network users that are authorized to push data up to the Private network. The enclave boundary protection system might also perform content review of the data submitted for release. Protection approaches that use Authenticated Parties are discussed below.

User and document authentication can be achieved with the Digital Signature and FORTEZZA technologies. Guards can check data packets for Digital Signatures or user identification and authentication and based on this information can accept or deny traffic from entering the enclave. The enclave boundary protection system cannot perform the functions of inspecting the contents of the message or verify the digital signature if the message is encrypted. Messages must be able to be decrypted prior to processing through the guard so that the guard will be able to perform filtering on the contents of the messages.

## Digital Signature

The digital signature, which is the result of encrypting a document using the private key of the signer, can be applied to spreadsheets, Word documents, e-mail messages, PDF files, and others. A digital signature is a string of numbers that is the representation of the document. Utilizing a digital signature ensures that the contents of a document cannot be altered; to do so would invalidate the signature. A digital signature is unique to both the signer and the document; therefore, user and document authentication can be achieved. However, the signature cannot provide confidentiality to the data contents.

An important note to recognize is the difference between the digital signature and a digitized signature. A digitized signature is simply the visual form of a handwritten signature to an electronic image. A digitized signature can be forged, duplicated, and cannot be used to determine if information has been altered after signature.

## Hardware Tokens

Hardware tokens can be used to identify and authenticate users. Hardware tokens include One-Time Only Passwords, FORTEZZA, and Smart Cards (the latter two are addressed in more detail below). One-Time Only Passwords protect against unauthorized access by providing dynamic user authentication. A personal identification number along with a code that changes very frequently (as frequent as every 30-60 seconds) is requested from the user for identification and authentication. A guard will process this information to permit or deny access. By requiring two factors of authentication, greater protection is provided against unauthorized access than with the traditional fixed password.

## FORTEZZA

Fortezza is a registered trademark held by the National Security Agency (NSA) that is used to describe a family of security products that provide data integrity, originator authentication, non-repudiation, and confidentiality. Fortezza is an “open system” allowing for seamless integration with most data communication hardware platforms, operating systems, software application packages and computer network configurations and protocols. This technology uses a cryptographic device; a PC Card called the Fortezza Crypto Card. This card contains the user's unique cryptographic key material and related information, and executes the public key cryptologic algorithms. Some abilities of the Fortezza Card are that it enables users to encrypt, decrypt, archive data as well as generate digital signatures. The card uses the Secure Hash Algorithm, Digital Signature Standard, Digital Signature Algorithm, and the Key Exchange Algorithm. A guard can identify and authenticate the originator of a message based on a digital signature. However, a guard must be able to decrypt traffic in order to determine permissibility into an enclave. If a guard is unable to decrypt data then the information will be denied from passing through the guard and entering the enclave.

## Smart Cards

Smart Cards are another technological method in which users can be identified and authenticated. A Smart Card is a plastic card embedded with a computer chip that stores and exchanges data between users. Smart Cards provide tamper-proof storage of user and account identity and add to system security for exchanging data across any type of network. They can serve as a means for network system, application or file access as smart cards can be used to obtain access to a computer or even e-mail accounts. Insertion of the card or close proximity to an antenna is required in order to be able to “read” the information on the card using a smart card reader that can be attached to a computer. Users can be authenticated and granted access based on preset privileges. A Guard can authenticate and identify users and thus determine access privileges into an enclave based on the information provided from the smart card.

## Secure Socket Layer (SSL)

Secure Socket Layer (SSL) is a popular security protocol for implementing privacy and authentication between communicating applications. It is a transport layer security protocol, enabling the encryption and authentication of arbitrary applications. The protocol prevents eavesdropping, tampering with information, and forging of information sent over the Internet.

The SSL protocol includes a lower level protocol (called the SSL Record Protocol) that encapsulates higher-level security protocols. The SSL Handshake Protocol is one such encapsulated protocol. It allows communicating parties to authenticate one another, and to establish cryptographic algorithms and keys at the start of a communication session. For more information about SSL, please visit <http://welcome.to/ssl>. [5]

Connections using SSL have three properties:

- The communication is private. The initial handshake uses public key cryptography to define a secret key. The secret key is then used with symmetric cryptography to encrypt all communications.

- Clients and servers can authenticate one another during the handshake using public key cryptography.
- The entire communication is protected against tampering or insertion of data. Each datagram has a Message Authentication Code that is a keyed hash value.

The SSL protocol can be used for network access between clients on the private side and servers on the public side. By checking a server's identity, confidence is obtained that the server is trusted to some degree. A policy requiring that SSL be used for all network access between private and public networks would effectively permit access only to those servers on the public side that have the ability to authenticate using SSL. However, the goal should not only be authentication. Rather, the goal should be access control, with authentication being a means to implement access control. This is accomplished by maintaining a list of public servers and directories that, once authenticated, can be accessed by private clients. That access list is best maintained by an enclave boundary protection system such as a Guard.

### 6.3.5.2 Confidentiality and Integrity

Confidentiality and Integrity can be assured through the following technologies, FORTEZZA, COTS Encryption, Audit Logs, and the Operating System.

#### FORTEZZA

In addition to the authentication and identification features of FORTEZZA, the cryptographic features of the "FORTEZZA Crypto Card" are employed to offer confidentiality and integrity. The integrity protection is provided primarily when data served from a server or client is key hashed (via the Secure Hash Algorithm FIPS PUB 180). [6] Confidentiality is accomplished with pre-encryption of the data to be served from the server, and the encryption/decryption of all data passed from a server to a client, and from a client to a server (via the Key Exchange Algorithm and SKIPJACK Algorithm FIPS PUB 185). [7] These cryptographic features also include digital signature capabilities, as well as the associated key and certificate management infrastructure support. FORTEZZA encryption and decryption functions include:

- Interface to and function with any Government-certified FORTEZZA Cryptographic Card for encryption and decryption.
- Not corrupt the integrity of a file's data content.
- Ensure that the resultant decrypted file retains the original file's attributes (e.g. if the original file was read-only, then when that file is decrypted, after being encrypted, it shall retain the read-only attribute).
- Be able to encrypt/decrypt files of all types.
- Inform the user if the encryption/decryption process succeeded or failed.
- Verify that any signature on the certificate is valid (based on the public key from the issuer's certificate).
- Allow the originator to select the type of protection to be applied to the message: signed-only, encrypted-only, or signed and encrypted.

## Commercial-Off-The-Shelf Encryption

Some guard products incorporate COTS encryption algorithms, such as triple Data Encryption Standard (DES). While these algorithms are not suitable to protect classified information, they may be used to segregate communities of interest in a protected environment. For example, two users with different privileges at the same classification level may use a commercial encryption algorithm to logically and reliably segregate their traffic. Other organizations that do not possess classified traffic, but rather sensitive traffic, may allow commercial algorithms to provide data confidentiality. In either scenario, commercial encryption may be used on the enclave side of the guard to provide logical data separation.

## Audit Logs

Audit logs maintain a record of system activity both by system and application processes and by user activity of systems and applications. In conjunction with appropriate tools and procedures, audit logs can assist in detecting security violations, performance problems, and flaws in applications and ensure data integrity. A computer system may have several audit trails, each devoted to a particular type of activity. Auditing is a review and analysis of management, operational, and technical controls. The auditor can obtain valuable information about activity on a computer system from the audit trail. Audit trails improve the accountability and integrity of the computer system. For example, audits can be used in concert with access controls to identify and provide information about users suspected of improper modification of data (e.g., introducing errors into a database). An audit trail may record “before” and “after” versions of records. (Depending upon the size of the file and the capabilities of the audit logging tools, this may be very resource-intensive.) Comparisons can then be made between the actual changes made to records and what was expected. This can help management determine if errors were made by the user, by the system or application software, or by some other source.

## Operating System

A guard cannot provide any degree of assurance if it is installed on an operating system with well-known vulnerabilities. In order to be effective, guard software must be developed on a trusted operating platform. Additionally, the guard software must make effective use of the security mechanisms and services offered by the operating system. Part of the guard development process should be documenting how the guard utilizes the operating system in an effective manner. Guards built on insecure operating systems should not be considered.

The operation and security level of a guard is dependent upon the operating system. The platform must be a trusted operating system with high-level security mechanisms. Hackers that become frustrated while trying to penetrate the guard will try to attack the underlying operating system in hopes of gaining access into the enclave. The operating system must have segmentation of processes in order to minimize the risk from hacker attempts. Segmentation of processes is the separation of system calls at the operating system level. This segmentation allows applications to use restricted portions of the operating system and denies user ability to penetrate different security levels. That is, a separate login and password is required for different command levels of the operating system. Usually, each security level of the operating system will have a limited command set in compliance with the security policy of the operating system. The system administrator should therefore hold a clearance that is at least equal to that of the highest network connected to the guard.

In an MLS environment, the strength of some guards rest within the user workstations and the gateways. Each user workstation and gateway must be installed with a trusted operating system. Guards trust users to make decisions on the classification/sensitivity of information. The trusted operating systems control access to information displayed on a user workstation and control the movement of information out of the multi-level network (MLN). The MLN must use a trusted operating system, defined as an operating system accredited to maintain the trust between sensitive information and the authorized users. In the MLN architecture, an authentication server controlling user logins and monitoring network system activity enhances this service.

### 6.3.5.3 Processing, Filtering, and Blocking Technologies

Protection approaches that logically fit within this category utilize various processing, filtering, and data blocking techniques in an attempt to provide data sanitization or separation between private network data/users and public network data/users. Data originating from the private network is implicitly labeled as private data though it may be asserted to be data of a lower sensitivity level by a private network user. Enclave boundary protection devices such as a Guard may perform automated processing and filtering techniques, and if such tests are successfully passed, the data is actually regraded by automated means. In the reverse direction, such approaches often incorporate data blocking techniques, (typically in firewalls but also in guards) to regulate the transfer of data from public network users to private network users. Use of certain protocols may be blocked and/or data may be processed or filtered in an attempt to eliminate or identify viruses and other malicious code transfers.

Information passed between public and private networks may be encoded as binary information in some applications, e.g., imagery, the size of the piece of information to be processed may be very large. The guard will have to reconstruct the entire message from multiple packets, which requires large working memory space. Then the guard must pass the information through filtering and processing rules. With large files, this may take a nontrivial amount of time. If any of the imagery files are time sensitive, i.e., used as part of a training exercise that requires commands to be issued based on the imagery files, the guard may add delay that degrades the usability of the information.

It must be emphasized that data transfer between private and public networks involve risks, and one must take steps to mitigate risk. Processing, filtering, and blocking techniques involve inexact attempts to filter private data from outgoing transmission through content checking against a pre-defined list of prohibited strings. Scanning and detecting virus-infected executables, and blocking executables are also conducted. Since there are an almost infinite number of possible executables, and malicious ones can be detected only through prior knowledge of their existence, the problem of detecting “maliciousness” in an arbitrary executable is not computable. This is exacerbated by the fact that there exist many executables that users wish to allow to cross the network boundary (e.g., Java applets, Active X controls, JavaScript, Word macros) and that they would therefore not wish to filter out or block. Only by performing a detailed risk management tradeoff analysis wherein operational needs are weighed against security concerns can these issues be resolved.

Protection approaches that use processing, filtering, and blocking technologies rely upon processing to allow the flow of information between two networks while attempting to detect and block the leakage of Classified data and attacks. Such approaches include Access Control Lists,

Malicious Code Detection, Content Checking, Application/Attachment Checking, and Public to Private Replication. These approaches are discussed in the following subsections.

## Access Control Lists

Access Control Lists give the ability to selectively allow users access to information. Access Control Lists are lists that identify which users are permitted access to secure files, databases, programs, and administrative power. Discretionary Access Control is used to restrict access to a file. Only those users specified by the owner of the file are granted access permission to that file. Mandatory Access Control occurs when the security policy is dictated by the system and not by the object owner. Before access can be permitted or denied, identification and authentication of the user must be available. Guards use the identity and authentication presented by the user to determine if an ACL applies to that user. For example, if an ACL requires authentication via digital signature then permission will immediately be denied to all users that do not authenticate with a digital signature. When a user authenticates with a digital signature, access permission will be granted if that user is on that ACL.

## Malicious Code Detection

Although not a part of the guard itself, malicious code detection is integral to providing the high assurance level associated with guards. Attachments that are opened by the guard must be sent to the malicious code detector to scan for known macro viruses or other malicious code. Files that are reassembled must also be scanned for known malicious code. The high assurance that can be provided by a guard can easily be undermined if the guard is allowed to pass information containing malicious code.

## Content Checking

Content checking service scans internal and external electronic mail to detect and remove content security threats. Dirty word search filters, which are configurable, may be applied to search for certain words and send rejection messages back to the originators' system. A Dirty word search scans messages for certain security-sensitive words, as defined by a word list. The content checking feature can be adequately defined, developed and verified to evaluate the contents of the data to be transferred through the guard to ensure that no information at a sensitive level is transferred to a lower level system.

## Application/Attachment Checking

Part of the application layer assurance offered by guards is application checking. This mechanism protects against attachments possessing improper file extensions. For example, the security policy for the organization may allow Microsoft Word attachments to pass through its mail guard. However, simply inspecting the file extension to verify that it is ".doc," is not enough to assure that the file is actually a Word file. The guard must launch its version of Microsoft Word and attempt to actually open the file. If the file cannot be opened, it either has errors, or is mislabeled and should not be allowed to pass through the guard. If the file can be opened, it should be passed to a gateway malicious code checker to check for Macro viruses. If no Macro viruses are found and it message passes all other content checking filters, the attachment may be allowed to pass through the guard.



## Public to Private Replication

Public to private replication allows users on a private network to receive data that originates on a public network, without having to explicitly request that the data be sent from the public servers. Replication can be used for network access by pushing data from a public network to a private network. It can give the private network any application that passes messages from one host to another. The primary security property of replication is the prevention of data flows from a private to a public network.

A common example of this technology is a database replication. If a node on a private network requires access to a database on a public server, the database can be duplicated on another server that is reachable by the private network. The guard controls the flow of information between the replicated database and the private node. The private node may only have read privileges to the database, and not be able to write, depending on the security policy for the private network. The ability to write to the database would be dependent upon the guards' private network. The ability to write to the database would be dependent on the guards' ability to reliably downgrade information. Other examples of replication are File Transfer Protocol (FTP), electronic mail, and Web Push protocols.

Replication does not reduce the potential risk that data replicated into the private network may be hostile executable code. In order to mitigate this risk a guard would have to be implemented so that data can be first replicated in this network guard. The guard inspects the data for potentially hostile code and ensures the data passes this inspection before being forwarded into a private network.

To prevent data leakage from private networks to a public network, replication does not allow a direct back channel to send message acknowledgments from a private network to the public network, to do so would allow a large covert channel. The replication acts as an intermediary, sending acknowledgments to the public sender, and receiving acknowledgments from the private recipient. The public sender cannot determine with precision the timing of the acknowledgments sent from the private side. Hence, the intermediate buffer within the replication process reduces the bandwidth of the back channel. This disconnects any direct communication from private networks to a public network.

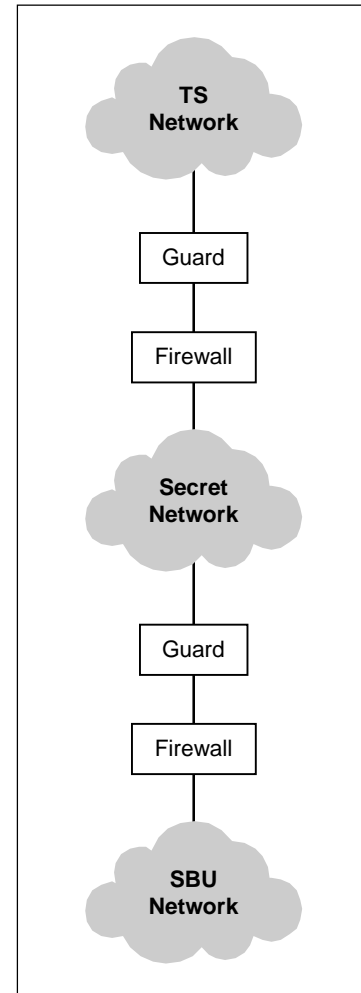
### 6.3.5.4 Cascading

Cascading occurs when two or more guards are used to connect three different networks containing information of three or more different levels. For example, if a TS (Top-Secret) and SECRET network establish an agreement and a connection and the Secret network has a pre-existing connection to a SBU network, the possibility exists for a path between the TS and SBU network. Please refer to Figure 6.3-3 below. The security policy for each guard needs to be examined to determine if a possible connection between the TS and the SBU network exists. Possible methods to reduce the risk associated with cascading are to allow different services through the two guards or restrict each user to interact with a single guard. When establishing a connection between two different networks using a guard, the connections each network have to other networks needs to be considered.

## 6.3.6 Selection Criteria

When selecting a Guard, the following should be taken into consideration.

- The guard should send and receive electronic mail between the private network and the public network.
- The guard should conform to standards used in the wider community.
- The guard should allow users to send and receive attachments in both directions.
- The guard should provide a user-friendly and seamless e-mail capability that passes messages with transit times comparable to those of a commercial electronic Message Transfer Agent (MTA).
- The guard should run on a trusted platform.
- The guard should only permit e-mail (Simple Mail Transfer Protocol (SMTP)) protocols to pass through the guard.
- The guard should allow only authorized users to send and/or receive a message by performing access control on both the source and destination addresses of the message.
- The guard should prevent message flow directly between the high side WAN and the guard in either direction.
- The guard should allow only a properly labeled message to pass from the private level to the public level --each message must include a classification label.
- The guard should ensure that the security level of a message subsumes (is equal to or greater than) the security level of its attachment(s).
- The guard should protect against unauthorized disclosure of information from a private network.
- The guard should provide safeguards to protect the private side from attacks (including penetration, malicious code, and denial of service) from the public side.
- The guard should allow word or phrase search.
- The guard should support user digital signatures and encryption applications.
- The guard should support a digital signature or encryption capability.
- The guard should audit all security-related functions.
- The guard should provide an access control mechanism to limit access to the guard's controls and provide separate roles for the security administration, system operator, and mail administration functions.



**Figure 6.3-3 Cascading Protection**

- The guard should provide rules-based sanitization (i.e., message content modification) of fixed format messages from high levels through low levels.
- The guard should ensure that only allowed data is distributed.
- The guard should validate the proper message construction, including configurable verification of message content.
- The guard should provide secure bridge for passing messages between networks of differing levels of security.
- The guard should downgrade high-level data from designated communications channels according to validated rules.
- The guard should verify that the data meets a set of rigorously controlled criteria.
- The guard should prevent disclosure or release data to unauthorized consumers.
- The guard should communicate with only specified hosts on the public networks.
- The guard should prevent workstations to be use as a pass-through or gateway device from the public sides for any communications, including mail.

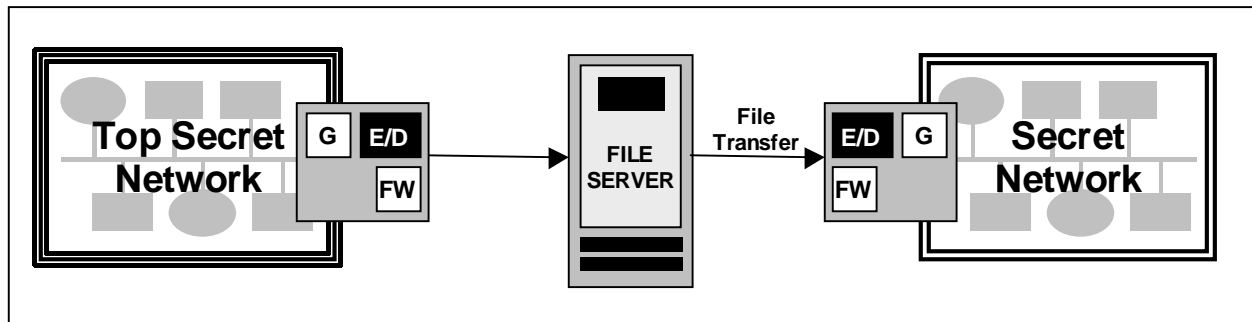
## 6.3.7 Framework Guidance

### 6.3.7.1 Case 1: Case 1: File Transfers from a Top Secret to a Secret Network

This Case study represents the situation when a user on a Secret Network must obtain files from a user on a Top Secret Network. Major risks are involved when connecting differing LANs. So when data files are to be transferred between networks of differing classification levels, the requirement arises for a guard that is able to recognize the File Transfer Protocol (FTP). Please refer to the Internet Engineering Task Force RFC 959 for additional information about the file transfer protocol, <http://www.ietf.org/rfc/rfc0959.txt?number=959> [8] The function of a guard is to try and permit communication between different classification boundaries while preventing the leakage of sensitive information. Included with the risks of connecting networks of differing classifications is the accidental or the malicious release of data from one network to another. Therefore when files must be transferred from a Top Secret Network to a Secret Network a guard can ensure that only permissible files are released. To be capable of this function a guard should be able to process files regardless of type (GIF-Graphic interchange format, Moving Pictures Expert Group file format (MPG), HTML, etc). The file will then be subject to review by the established application checking policy to scan the contents and verify the sensitivity level. The guard will then downgrade files to allow releasability of the file to a lower sensitivity level user. Downgrading occurs only if the content of a file meets the requirements of the sensitivity level of the network for which the data is being delivered. Downgrading is the change of a classification label to a lower level without changing the contents of the data.

Additionally, limits must be placed as to which users have permission to release files from the Top Secret Network and which users on the Secret side have permission to obtain these files. The originator of a file will have permission granted through an ACL kept by the guard to re-

lease files to the lower level, Secret, network. In return, the recipient must also have permission granted to access files that were released from the Top Secret network. Owners of data must be able to restrict access to their data and the system must also be able to deny access. Discretionary Access Control (DAC), is the access control mechanism that allows the owners of files to grant or deny access to users. The owner of a file can also specify an access control list to assign access permission to additional users or groups. Mandatory Access Control (MAC), is a system-enforced access control mechanism that uses clearances and sensitivity labels to enforce security policy. MAC associates information requested from a user with the users accessible security level. If data is classified as top secret the owner of the information cannot make it available to users who do not have access to top secret data. When access is restricted authentication and authorization policies must be in place. Authentication verifies the claimed identity of users from a pre-existing label. Authorization is the determination of privileges a user has in order to grant permission for access of requested information. Authentication and authorization must be performed for all users requesting sensitive files from a user as depicted in Figure 6.3-4. Files may be stored on a server making the files available to users on the Secret networks that have permission to access the files. The server that allows the release of files shall be a COTS product that receives files and places them in a directory so that they will be accessible to authorized users. A guard must also be configurable to allow changes to be made to a database. Changes made to the master database of downgraded data shall be applied to replicated databases in near real-time.



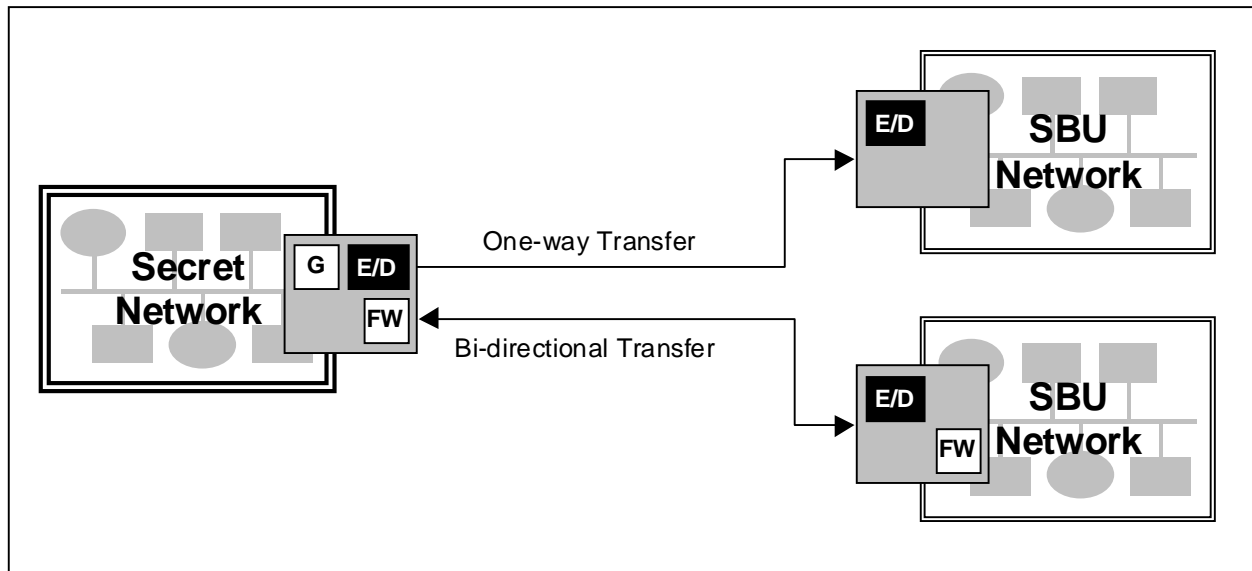
**Figure 6.3-4 File Transfers**

In keeping with the established releasability policy for file transfers, the guard will release the data to the lower level (secret) network based on the match of the content label and the security attributes of the recipient. The releasability policy the guard will follow shall adhere to the following.

- The guard shall only allow a very small set of users on the Top Secret side to release files.
- The guard shall maintain an access control list of these users and check the list every time a file is submitted for release.
- Only files of a specific format (plain text or HTML) shall be releasable.
- Strict audit logs shall be kept on the guard of all released files.
- Released files shall be scanned for content.
- Images contained within a file shall be reviewed.
- Authentication of all files (for example, digital signatures).

### 6.3.7.2 Case 2: Releasability from Secret to Sensitive But Unclassified Networks

When opening communication channels between secret and SBU (Sensitive but Unclassified) networks the determination shall be made as to whether a bi-directional flow of information through a guard will be allowed. Guards differ in that some support only one-way transfers of information while others support a bi-directional flow of information. Releasing information from a Secret to a SBU network can be done through email transmissions. Therefore, a mail guard is required as shown in Figure 6.3-5 and can be coupled with a firewall to further enhance the security measures taken to protect the Secret enclave.

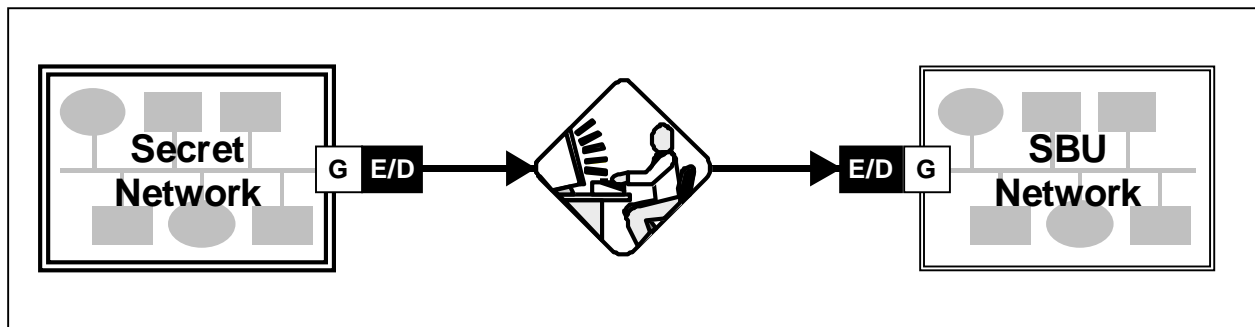


**Figure 6.3-5 Secret to SBU Releasability**

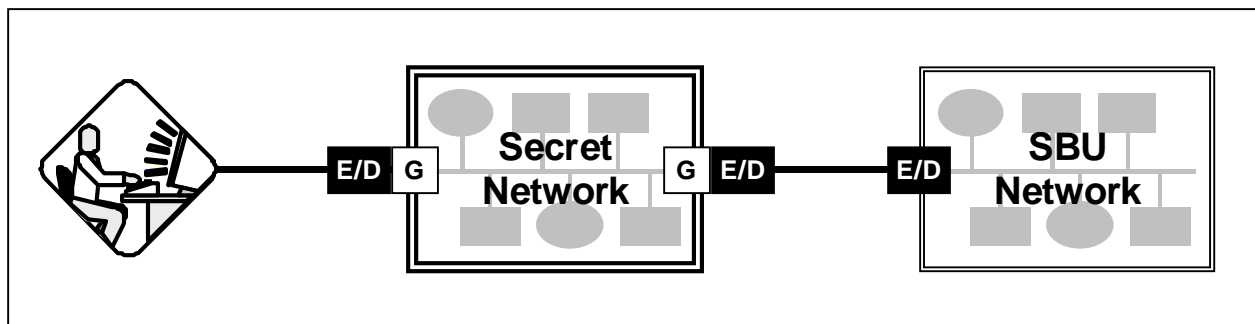
The mail guard enforces the policy for release of messages from the Secret network. This policy can include:

- Content filtering/dirty word search,
- Malicious code checking,
- Message format check,
- Envelope filtering to determine if a sender and receiver are permitted to send and receive messages,
- Authentication (for example, cryptographic digital signatures),
- Message journaling/logging and,
- The allowance/disallowance of attachments
- Review of attachment
- The allowance/disallowance of mail receipts,
- The allowance/disallowance of sending blind carbon copies of messages,
- Maintain and review audit logs of all mail message transfers for questionable actions.

While the goal is to have a guard that has full functionality and is able to automatically review all information, a human reviewer may also be placed to review messages prior to the guard receiving and reviewing messages. A person can manually review messages by being placed between the guards of two separate networks, as shown in Figure 6-3-6. Or, as shown in Figure 6.3-7, a human reviewer can review information prior to the guard for verification that the sensitivity level of the information can be released to the SBU network.



**Figure 6.3-6 Human Reviewer-Man in the Middle**



**Figure 6.3-7 Releasability Human Verification**

The human reviewer has the release authority over a message, whether to allow or reject the sending of the message. The established security policy may require that all messages are reviewed or only rejected messages are reviewed, or perhaps messages may not need to be manually approved. The functionality goal of a guard is to allow a fully automated review process. A process without a human reviewer must have fully automated guards that are able to check content, check attachments to email messages, have a configurable security filter, perform dirty word searches, and have imagery processing capabilities. Dirty word searches are looking for words or codes that could be used to disclose sensitive information.

Encrypted messages must be able to be decrypted prior to processing through the guard, allowing the message to be released. Guards with decryption capability (which may be through embedded FORTEZZA cards) will decrypt a copy of a message and upon release approval pass the original message to the recipient and discard the decrypted copy. If a message is unable to be decrypted then the guard must reject that message. A rejection notice policy shall be established to address the handling of message rejection notices. The rejection notice policy may have notices sent only to the mail administrator of the Secret network or can also allow rejection notices to be sent to the user. A policy shall also be established as to the allowance of mail receipts.

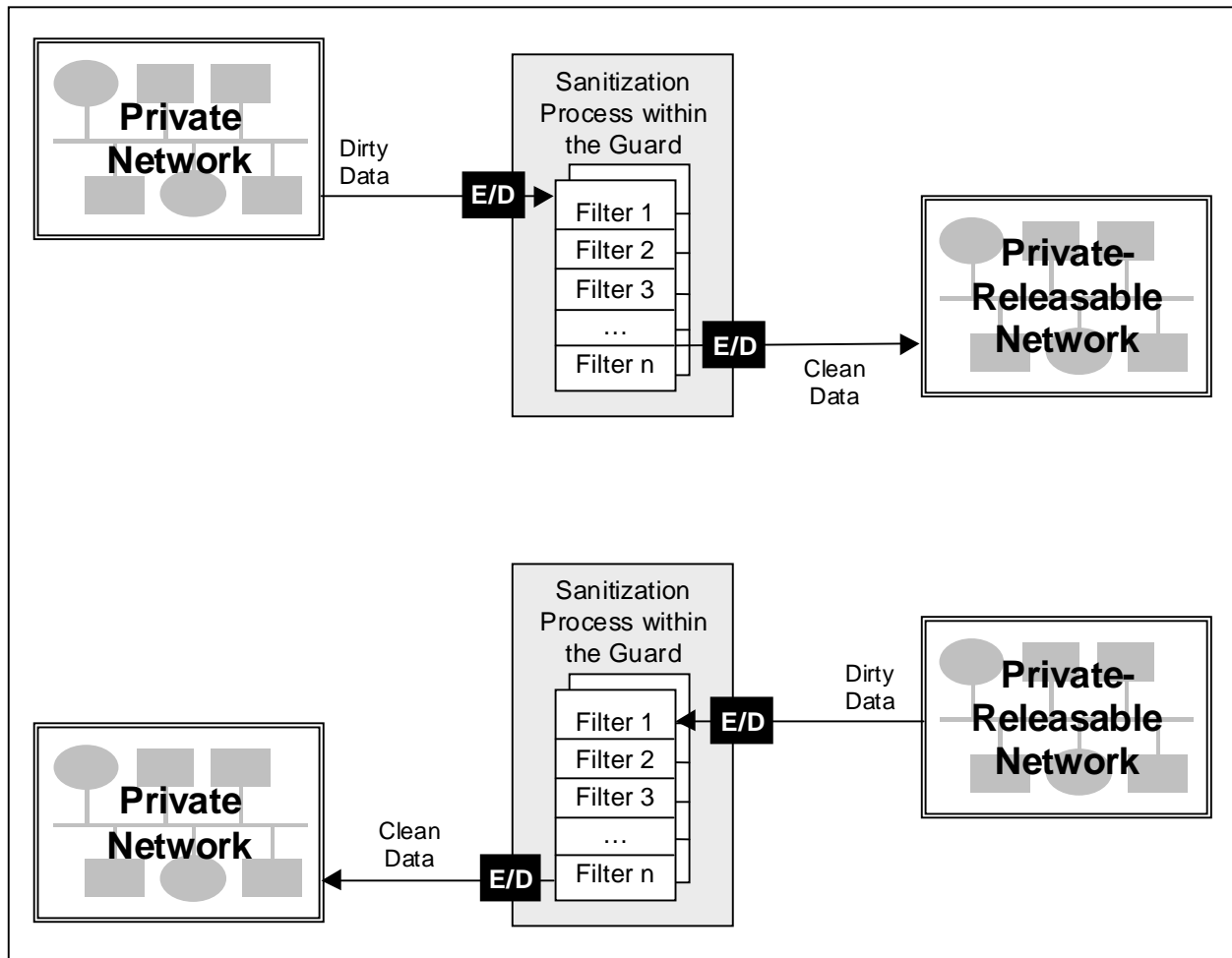
Confirmation that recipients have received an email can be equally important as the security measures taken to protect the information contained within the email. Mail receipts however, cannot always be relied upon, as some email servers will not allow receipts out of their own email system. Therefore, when sending secret information to SBU users, rules must be established as to the allowance of return receipts. Automatic return receipts may not be part of the security policy. However, once a recipient verifies that the appropriate message was received, a signed receipt can be generated and sent to the guard for filtering and then forwarded to the originator. In place of return receipts, servers capable of providing automatic tracking capabilities can be used to confirm document receipt.

Remote Access capabilities pose a risk as a backdoor mechanism to gain access into a network. Therefore, for this scenario, the Guard security mechanism would be most effective if coupled with a firewall. A firewall will protect the LAN from Internet or modem attacks by blocking direct access. Besides maintaining network access controls, the firewall will also maintain extensive audit records detailing successful and unsuccessful attempts to access the system. Once connected and authenticated, a dial-in user then has the same Internet services as local users. Internet connectivity is an inherent risk as it opens up channels of additional risk when connecting Secret networks to SBU networks. Therefore, a guard must be able to recognize Web-based protocols (i.e., HTTP) to mitigate risk for access into the networks.

Another important means of communicating for business is real-time messaging. Therefore, guards should be able to support real-time/instant messaging. When communicating by means of real-time messaging, messages should be ensured against corruption, tampering, recording, and non-playback.

### 6.3.7.3 Case 3: Private to Private-Releasable

This case involves information that is passed from Private Networks to Private-Releasable Networks. A Private-Releasable network is a peer of a Private Network in that they are both of the same security level. However, each network contains information sensitive to that network/organization. A bi-directional guard may be more applicable to this case situation in order to allow information to flow between the networks regardless of the network of origination.

**Figure 6.3-8 Sanitization Process**

Information flowing from a Private network to a Private-Releasable network must pass through the sanitization process as shown in Figure 6.3-8. When sensitive data is passed between networks of differing sensitivity levels sanitization is necessary. Sanitization is the changing of the contents in order to meet the requirements of the information sensitivity level of the network for which the information is destined. Information that is passed from a network of one sensitivity level to a network of another sensitivity level is considered 'dirty' upon release from a private network. Information is termed to be dirty because users of a differing sensitivity level are not allowed access to this information. After being processed and sanitized through the guard the information is termed to be 'clean' and therefore accessible by those users holding privileges of a differing sensitivity level.



The releasability policy for data being transferred between private and private-releasable networks should be similar to that of Case 1 above. Only files of a specific format (plain text or HTML) shall be releasable and strict audit logs shall be kept on the guards of all files transferred between both networks. Access Control Lists may also be maintained to allow only authorized users to transfer or receive files.

## 6.3.8 Technology Gaps

### 6.3.8.1 High Volume of Binary Data

Some applications require that information be passed in a binary representation. Examples of these applications are voice, imagery, and video. Binary data is more difficult to perform content checking on and to pass through filter rules. Guard technology needs to become faster to allow large amounts of binary files, and streaming binary information to pass through the high assurance mechanisms to which other information is subject.

### 6.3.8.2 Quality of Service

Quality of service (QoS) is being deployed in networks to support real time applications, such as voice and video as well as for other applications that may have strict latency requirements. There are several different approaches to supporting QoS in IP networks. While currently, there are multiple approaches for providing QoS in an IP network, the guard that is implemented must support the QoS strategy for the organization.

Guards must support the quality of service mechanisms provided by the network. All incoming traffic is passed through the guard. If the QoS mechanism is not supported by the guard, end-to-end QoS that is required by the application cannot be supported.

### 6.3.8.3 High Speed across Optical and Other Networks

Currently, most guards are designed to work in IP networks. However, there are many different types of networks that could make use of guard technology including all optical networks and Asynchronous Transfer Mode (ATM) networks. These networks typically operate at speeds in excess of those of IP networks. In addition to adding the proper interface to the guard, the filtering mechanisms within the guard must be capable of the speeds on the optical network. Additionally optical and ATM networks are more sensitive to delays. If the guard is not capable of supporting the bandwidth requirements of a connection, communications through the guard may be degraded to the point where further connections cannot be accepted.

### 6.3.8.4 HyperText Markup Language (HTML) Browsing

Today's network environment makes use of HTML traffic for a variety of applications. Having a guard that supported HTML browsing for Internet or internal HTML would greatly increase the functionality of organizations.

In order to support HTML, a guard would have to allow requests, (i.e., Domain Name Server (DNS) queries, requests for web pages) to pass through the guard. When the response returns, the guard must intercept the message and perform its checking before it is allowed to pass back to the user. All this must happen in real time to allow for human interaction and viewing behind the guard.

# References

1. Computer Advisory Incident Capability. Department of Energy. 6 June 2000  
<<http://www.ciac.org/ciac/ToolsTestEmail.html>>.
2. CERT® Coordination Center. 17 July 2000 <[www.cert.org](http://www.cert.org)>.
3. Software Engineering Management Practices. Carnegie Mellon Software Engineering Institute. 18 July 2000. 12 June 2000  
<<http://www.sei.cmu.edu/managing/managing.html>>.
4. Trusted Product Evaluation Program. 12 June 2000.  
<<http://www.radium.ncsc.mil/tpep/tpep.html>>.
5. Lashley Brian and Andrzej Tarski. SSL <<http://welcome.to/ssl>>.
6. Federal Information Processing Standards Publications (FIPs) Pub 180. Secure Hash Standard 17 Apr 96 <<http://www.itl.nist.gov/fipspubs/by-num.htm>>.
7. Federal Information Processing Standards Publications (FIPs) 185. Escrowed Encryption Standard. 09 Feb 94 <<http://www.itl.nist.gov/fipspubs/by-num.htm>>.
8. Postal, J. and J. Reynolds. "File Transfer Protocol (FTP)." RFC 959, ISI, 1985 October.  
<<http://www.ietf.org/rfc/rfc0959.txt?number=959>>.

# Additional References

- a. Computer Advisory Incident Capability. Department of Energy. 6 June 2000  
<<http://ciac.llnl.gov/ciac/bulletins/I-005c.shtml>>. Enter at <<http://ciac.llnl.gov>>, then navigate to: <<http://ciac.llnl.gov/ciac/bulletins/I-005c.shtml>>.
- b. Digital Signature Trust Co. 3 July 2000 <<http://www.digsigtrust.com/>>.
- c. F-Secure Corporation. "Introducing IPSec White Paper". March 1999. Online posting. 21 June 2000  
<<http://www.Europe.F-Secure.com/solutions/white-papers/ipsec0399.htm>>.
- d. National Institute of Standards and Technology (NIST) Fips 186. FACT SHEET ON DIGITAL SIGNATURE STANDARD. Online posting May 1994. 3 July 2000  
<[http://www.nist.gov/public\\_affairs/releases/digsigst.htm](http://www.nist.gov/public_affairs/releases/digsigst.htm)>.
- e. NetworkWorldFusion News. 20 June 2000  
<<http://www.nwfusion.com/news/tech/0906tech.html>>.
- f. Stronghold Webserver Administration Guide Chapter 6 SSL Authentication and Encryption. 22 June 2000 <[http://mclean2.his.com/docs/Administration\\_Guide/SSL.html](http://mclean2.his.com/docs/Administration_Guide/SSL.html)>.
- g. Stronghold Webserver Administration Guide Chapter 6 SSL Authentication and Encryption. 22 June 2000 <<http://developer.netscape.com/docs/manuals/security/sslin/contents.htm>>.

**UNCLASSIFIED**

Guards

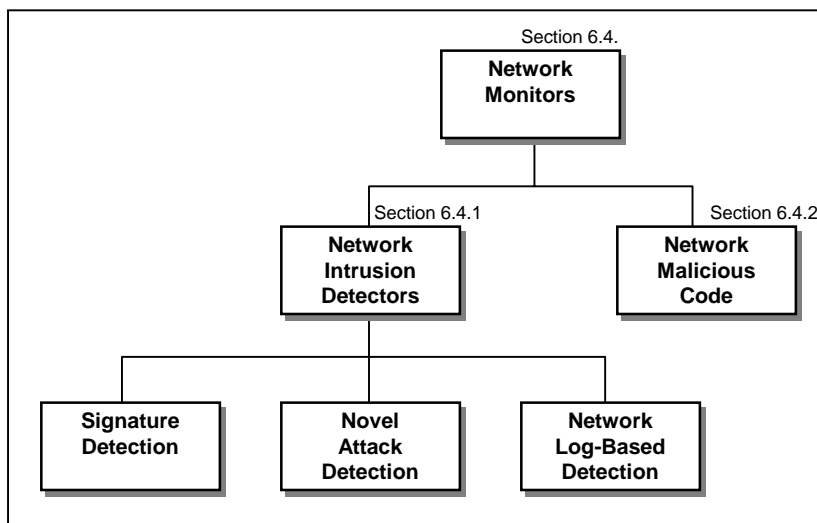
IATF Release 3.0—September 2000

- a. The Source of JAVA™ Technology. Smart Card Overview. 5 July 2000  
<<http://www.java.sun.com/products/javacard/smartcards.html>>.
- b. Smart Card Basics.com. 5 July 2000 <<http://www.smartcardbasics.com/security.html>>.
- c. Hulme, George V. "Secure Document Delivery Gains Favor." InformationWeek. 17 July, 2000.

## 6.4 Network Monitoring Within Enclave Boundaries and External Connections

A fundamental tenet of the defense in-depth strategy is to prevent cyber attacks from penetrating networks and to detect and respond effectively to mitigate the effects of attacks that do. As discussed above, an integral aspect of the defense in-depth strategy embraced by this framework is enclave boundary protection, which often takes the form of firewalls and Virtual Private Network (VPNs). While these technologies offer perimeter and access controls, “authorized” internal and remote users can attempt probing, misuse, and malicious activities within an enclave. Firewalls don’t monitor authorized users’ actions, nor do they address internal (insider) threats. Firewalls must allow some degree of access, which may open the door for external vulnerability probing and the potential for attacks.

Detect and respond capabilities are complex structures that run the gamut of intrusion and attack detection, characterization, and response. The various Detection aspects of Detect and Respond are actually measurement services. Intrusion detection, network scanning, and the like are measurement functions that determine the effectiveness of the deployed protection systems and procedures on a continuous and/or periodic basis. In themselves, Detect capabilities are not Protection measures. The respond aspect can initiate changes to existing protection systems (e.g., configuration change on a firewall to block an attacker’s Internet Protocol (IP) address) or the deployment of additional protection measures (e.g., placement of another firewall appliance). The local environments (within enclaves) are the logical location for network-based sensors. This section addresses sensors that operate in near real time. Specific network monitor technologies addressed in the Framework are shown in Figure 6.4-1. Section 6.5, Network Scanners Within



**Figure 6.4-1 Breakdown of Network Monitor Technologies**

Enclave Boundaries, addresses sensors that typically operate off-line. Section 7.2 (Host-based Detect and Respond Capabilities within Computing Environments) provides similar guidance for host-based sensors.

Local environments have the option to implement as much or as little above the sensors as they believe is prudent, obtaining services and support from the infrastructure as necessary. Section 8.2 of the framework provides an in-depth discussion of the various Detect and Respond processes and functions in the context of a supporting IA infrastructure capability. It also offers

guidance on technologies for processes beyond the sensors, but recognizes that they may be implemented at any level in a network hierarchy (including a local enclave environment).

Network monitors, including network intrusion detection and network malicious code detector technology areas, are covered in this section. The section provides an overview of each relevant technology, general considerations for their use, rationale for selecting available features, deployment considerations, and a perspective on how these technologies are typically bundled into products. The section concludes with sources for additional information and a list of references used in developing this guidance.

## 6.4.1 Network Intrusion Detection

The goal of an Intrusion Detection System (IDS) is to identify and potentially stop unauthorized use, misuse, and abuse of computer systems by both internal network users and external attackers in “near real time.” As this section of the framework addresses network-based monitoring, these discussions center on operations using network information. As discussed in Section 7.2 (Host-based Detect and Respond Capabilities within Computing Environments), similar structures and technologies are also available for performing comparable functions using host-based information.

### 6.4.1.1 Technology Overview

Normally, a dedicated computer is deployed for each network IDS on each network or network segment being monitored. A Network Interface Card (NIC) is placed into promiscuous mode, enabling the IDS software to watch all traffic passing from computer to computer on that particular network. The IDS software looks for signs of abuse (e.g., malformed packets, incorrect source or destination addresses, and particular key words).

A network based IDS forms its attack detection upon a comparison of parameters of the user’s session and the user’s commands to a rules-base of techniques used by attackers to penetrate a system. These techniques, referred to as “attack signatures,” are what network-based IDSs look for in the behavior of network traffic. An attack signature can be any pattern or sequence of patterns that constitute a known security violation. The patterns are monitored on the network data. The level of sophistication of an intrusion can range from single events, events that occur over time, and sequential events that comprise an intrusion.

## Detection Approaches

There are three basic technology approaches for performing network intrusion detection:

- *Signature detection approach* typically incorporates search engines that look to identify known intrusion or attack signatures.
- *Novel attack detection* is based on identifying abnormal network behavior that could be indicative of an intrusion.
- *Network log-based detection* monitors for attacks using audit logs of network components.

*Signature Detection Approach*—This approach utilizes traffic analysis to compare session data with a known database of popular attack signatures. These IDSs act like a “sniffer” of network

traffic on the network, caching network traffic for analysis. Typically, they do not introduce path delays while they are processing traffic and therefore do not impact network or application performance. Vendors refer to this operation as real time. Northcutt offers the perspective that “one of the great marketing lies in intrusion detection is ‘real time.’ What marketers mean by real time is that intrusion detection analysts are supposed to respond to beeps and alarms.” [“Network Intrusion Detection An Analyst’s Handbook,” by Stephen Northcutt, New Riders Publishing, 1999]

This technology examines the traffic against a predefined set of rules or attack signatures, typically using one of these techniques:

- Pattern expression or bytecode matching—the ability to determine regular behavior patterns to distinguish abnormal patterns, as well as determine if the traffic being monitored matches a predefined “attack signature.”
- Frequency or threshold crossing—the ability to establish a predefined threshold; if the threshold is exceeded, an intrusion is assumed.

There are two basic signature-based options: one, referred to as a *Static Signature IDS*, which utilizes uses a built-in attack signature base and a second, *Dynamic Signature IDS*, that relies on signature information that can be loaded dynamically into the IDS. Some product vendors provide routine updates of attack signatures. Some IDS tools give the customer the capability to customize attack signatures.

*Novel Attack Detection*—This relatively new detection strategy monitors Transmission Control Protocol (TCP) Dump data and attempts to filter out activities that are considered normal behavior. The genesis for this approach was to implement a sensor that would allow an analyst to evaluate large quantities of network information and select anomalous behavior. Unlike signature detection techniques, where the sensor has to have a priori knowledge of specific attack scripts, this technique relies on a screening by an analyst and can detect a variety of probes and attacks that other detection approaches miss. Initial versions dealt with packet header information only. Later versions capture the full packet content.

*Network Log-based Detection*—This detection technique focuses on the monitoring of audit logs from network devices. It has two major components. One is a catalog of audited events that are considered “bad” behavior. The catalog could include attack profiles, suspicious activity profiles, and activities defines as unacceptable. The second component is an audit trail analysis module. Audit trails come from a chronological record of activities on a system. The analysis module examines the monitored system’s audit trail for activity that matches activity in the catalog; when a match occurs, intrusive activity is assumed. Audit-based systems may also provide the ability to identify and track additional activity performed by an individual who is suspected of intrusive behavior.

## IDS Tuning Options

Typically, an IDS provides capabilities for selecting which attacks are being monitored. Depending on the specific implementation of an IDS, it is often possible to select which attacks will be monitored, what the response will be for each detected intrusion, specific source and/or destination addresses (to be monitored or excluded), and characterizations of the class (indication of the importance or severity) of each alarm. This capability is critical to optimize

the monitoring capability for an IDS. In this way it is possible to focus the sensor on specific events of interest, and the response that the IDS will have on the detection of events.

## Response Options

While the sensors detect and collect information about intrusions, it is the analyst that interprets the results. Some network IDS technologies offer automated response features to various alarms. In addition to logging the session and reporting as indicated below, some have the option to terminate the connection, shun an address that was the source of the detected intrusion, throttle the amount of traffic allowed through a port, or even close down a site's operation. In some cases, the IDS can accomplish these operations itself; in others, it works in conjunction with a network interface device (e.g., firewall, router, or gateway) to achieve the desired result.

## Reporting Mechanisms

When it detects a threat, a network IDS generally sends an alert to a centralized management console where alert information can be recorded and brought to the attention of an administrator. Some of the network IDS technologies offer additional reporting capabilities. Some can automatically send an e-mail message over the network to alert an operator to the alarm condition. Others can initiate a message to a pager.

### 6.4.1.2 General Considerations for Use

Network IDS technologies provide an important aspect of an enclave's defensive posture. Table 6.4-1 provides a synopsis of advantages and disadvantages of using network-based IDS technology.

**Table 6.4-1 Network-based IDS Considerations**

Advantages	Disadvantages
<p>Provides real-time measure of the adequacy of an infrastructure's network protection measures.</p> <p>Network-level sensors can monitor and detect network attacks (e.g., SYN flood and packet storm attacks).</p> <p>The insertion of a network-level sensor does not affect existing data sources from a performance and reliability standpoint.</p> <p>Well-placed network sensors are designed to provide an integrated, enterprise-wide view, at the management console, of any large scale attack.</p> <p>Operator expertise and training only for required for the single network IDS platform</p>	<p>Some network-based systems can infer from network traffic what is happening on hosts, yet they cannot tell the outcome of the commands executed on the host.</p> <p>Network-based monitoring and intrusion detection becomes more difficult on modern switched networks. Switched networks establish a network segment for each host; therefore, network-based sensors are reduced to monitoring a single host. Network switches that support a monitoring or scanning port can at least partially mitigate this issue.</p> <p>Network-based sensors cannot scan protocols or content if network traffic is encrypted.</p> <p>Must be used on each network segment because they are unable to see across routers and switches.</p> <p>Current network-based monitoring technologies cannot handle high-speed networks.</p>

The network-based IDS is typically deployed in the middle of a communications path between client and server and has access to data at all layers of communication. This process



allows this type of sensor to do extensive analysis for attack detection and provide detection in “near real time.” Since a network IDS runs on an independent computer, there is no impact on the performance of other network resources.

Today, network traffic is often encrypted through mechanisms such as VPNs. A network IDS simply watches information traversing a network and is typically not capable of decrypting the packets. In these cases, the encryption blinds it to any attacks that may occur. This type of sensor relies on passive protocol analysis causing it to “fail open.” This leaves the network available and vulnerable, and leaves the IDS itself open to potential compromise.

Throughput is another concern, if only one network IDS computer was to monitor an entire network, that one computer would have to be capable of scanning every single network packet. At modest levels (e.g., 50 Mb/s) most network IDSs can keep pace with the incoming stream of data. However, as network bandwidth increases and network loads reach higher rates (100 Mb/s and beyond), one or even several network IDS computer may not be able to keep up with the flow of traffic.

### 6.4.1.3 Important Features

When considering the selection of a network IDS, there are a number of features that should be considered. This section identifies important features for selection. The section that follows discusses rationale for the selection of these features.

#### ***Detection***

- Detection approach used by the network IDS.
- Does it perform packet fragmentation/re-assembly?
  - Which threshold adjustments can be made to the IDS?
- Which threshold adjustments can be made to the IDS?

#### ***Signatures***

- Number of events/signatures that can be stored.
- How often the signatures can be updated?
- Is the update static (Manual) or dynamic (Automated)?
- Are user-defined attack signatures allowed; if so, are the scripting tools easy to use?

#### ***Operations***

- Can it protect itself from unauthorized modifications?
- It recover from system crashes?

#### ***Response Options***

- Does it offer provisions for reconfiguring firewalls?
- Does it have session closing and reset capabilities?
- Does it have address blocking (shunning) capabilities?
- Can it execute program scripts on alarm?

### ***Reporting Options***

- Does it report in “real time” to a workstation?
- Can network and host-based IDS report to the same analyst console?
- Is the reporting interval configurable?
- Can it notify personnel using e-mail or pagers?
- Is the amount/ type of information reported to a management station configurable?

### ***Performance***

- Network compatibility.
- Number of packets that can be processed over an interval (packet size/bandwidth).
- Rate of False Positives (identification of a non-intrusive activity as intrusive).
- Rate of False Negatives (failure to identify an intrusive activity).

### ***Platform***

- Operating System.
- Type of platform required to host network IDS.
- Processing burden for anticipated network traffic load.

### ***Console Considerations***

- Operator Interface—Type of command and monitoring provisions available to an operator.
- Mark As Analyzed—Ability to clear or mark selected alarms that have been reviewed  
Drill Down—Ability to provide additional information for selected events.
- Correlation—Tools to correlate events based on source, destination, type.
- Report Generation—Ability to generate reports upon event detection and as periodic summary reports.

## **6.4.1.4 Rationale for Selecting Features**

Detect and respond capabilities exemplify the necessity of integrating operations and personnel considerations with the selection of technology solutions, consistent with the overall defense in depth philosophy. As indicated earlier, network monitoring does not itself offer protection from intrusions or attacks. It should really be considered instrumentation that monitors (and “measures”) the effectiveness of a network’s existing protection structures. It is up to operators (personnel and operations) to interpret the outputs of the IDS and initiate an appropriate response. If full time operators<sup>1</sup> are not available to interpret and formulate responses based on the IDS outputs, then IDS implementations will not typically add real value. In this case, it is likely that IDS deployments should not be considered. Otherwise, when considering features

---

<sup>1</sup> Ideally operators should be available on a 24x7 basis. The number of operators will depend on the traffic loads and anticipated numbers of incidents. It is not uncommon to experience hundreds of thousands of intrusion alerts per day, and each must be investigated to determine which, if any are serious threats.

for an IDS, there are a number of factors to be considered that are based on how the IDS is intended to be used, whether full or part-time operators will be available, and the skills of the operators to interpret the results.

## Detection

The type of detection mechanism is one primary consideration when selecting a network IDS technology. Another important consideration is the anticipated skills of the attacker. Signature-based detection, which is the traditional method used in network IDS technologies, typically lacks the ability to detect new (or modified) versions of attack strings. While many intrusions (typical from novices) use standard attack sequences (often downloaded from hacker bulleting boards), an accomplished adversary will have the capability to create new attacks or modify old attacks and thus thwart its traditional signature detection mechanisms. Anomaly and misuse detection approaches have greater flexibility for identifying new or modified attacks (since they monitor network usage or behavior). They are more complex to operate and not necessarily as responsive to traditional attack strings. These are also the only mechanisms currently available to monitor actions of otherwise authorized users for inadvertent or intentional misuse.

The ability of the various detection schemes to correctly identify intrusions is a fundamental consideration. The rate of false positives (alerts resulting from normal traffic) and false negatives (failure to identify a real intrusion attempt) should be considered. While the technologies are continually being refined for improved performance, there are inherent features that may limit performance (e.g., anomaly detectors have been known to generate significantly higher false positive indications).

As always, any decision is based on level of risk, anticipated performance, cost (for purchase, deployment and operation), and operational impact. The Framework recommends consideration for deployment of multiple attack detection schemes, ideally from different vendor sources. In this way, there is a greater likelihood of detection by at least one of the mechanisms deployed.

## Signatures

If a signature-based IDS is selected, it is desirable to have as many signatures as possible used for detection. However, there is usually an inverse relationship among the number of signatures, and the response time for possible detection, and the amount of traffic that can be monitored. Since the “lists” of possible attacks change frequently, it is strongly recommended that the IDS be capable of dynamically loading signatures. It is usually operationally more feasible and efficient if the downloading is handled on an enterprise (or at least site) basis. Most vendors that offer dynamic loading of signatures provide periodic updates to their signature base. While the update periods differ among vendors, a good rule of thumb is the more often the better. If operators have the skills to create custom signatures, then having the ability to support user-defined attacks is also desirable, particularly if custom attacks are found in one of your sites.

## Operations

It is desirable for the IDS to be easily configurable according to the security policies of the information system that is being monitored. Consideration should also be given to its ability to

adapt to changes in system and user behavior over time (e.g., new applications being installed, users changing from one activity to another, or new resources being available that cause changes in system resource usage patterns).

By its nature, IDS sensors are located where intrusions are anticipated. Thus, it is important that an adversary not be capable of modifying the IDS to render it ineffective. It is desirable that the IDS is able to perform self-monitoring, detect unauthorized modifications, and notify an attendant console. To simplify recovery of operations after an intrusion, it is also desirable that the IDS be able to recover from system crashes, either accidental or caused by malicious activity, and upon startup, be able to recover its previous state and resume its operation unaffected.

## Response Options

Many available solutions offer automated response options that seem on the surface to be very desirable. They infer that little or no human interaction is involved, as the devices can provide an immediate response. There is a serious pitfall to consider, however, before these options are deployed. First, it is not uncommon for a network IDS to find thousands (and possibly hundreds of thousands) of events daily, depending on where it is employed, characteristics of the “normal” network traffic load, and many other factors. Often, the number of false positives may be high, giving frequent rise to unwarranted indications of intrusions. Automated responses that terminate connections, shun addresses, throttle traffic, or actually shut down a facility can often cause severe denial of service threats to the network. It is strongly recommended that automated options not be used if there is a concern that they may cause denial of service on the networks they are trying to defend.

## Reporting Options

Most network-based IDSs report alarms to an operator console. (Please refer to the discussion of console features below.) The desirability of the level and frequency of reporting is based primarily on the availability and skills of the operators. Some network IDS technologies offer the option of paging or sending e-mail messages to notify personnel of alarms. While these sound desirable, they have the potential to give rise to operational issues. With an IDS detecting thousands of alarms a day, these features have the potential for overloading e-mail servers (creating a denial of service threat themselves), or paging operators extremely frequently at all times of the day and night. Most often, these features are not recommended.

## Performance

Network intrusion detection system performance varies due to the speed of the network, the amount of traffic, number of nodes being protected, the number of attack signatures employed, and the power of the platform on which the IDS resides. Identification (ID) systems may be overtaxed. On busy networks. However, multiple IDSs can be placed on a given segment to subdivide host protection, thereby increasing performance and overall protection. For instance, high-speed networks employing Asynchronous Transfer Mode (ATM), which uses packet fragmentation to improve efficiency over high bandwidth communications, do pose problems in terms of performance and response.

## Platform

A major issue for the selection of a network-based IDS is the type of computer skills (e.g., UNIX, NT) required for operators. They will likely need these skills to perform installation, configuration, adjustment, and maintenance. Since a network-based IDS usually is located on its own platform, the platform will have to be acquired and maintained, so it may be useful to select a technology that functions on the types of platforms used within the enterprise.

## Console Considerations

As discussed in Section 8.2 of the framework, the primary function of the console is to serve as an aid in the characterization and analysis of the many alarms that will be identified. Operators will have to not only identify alarms that were unwarranted, those that do not offer serious risks to the network, and those that do, as well as gain a first order understanding of the source of and impact of possible attacks

*Operator Interface* – The type of interface that is operationally desired tends to be driven directly by operator preference. Novices typically prefer a Graphical User Interface (GUI) with intuitive operations, pull down screens, and substantial aids available. Skilled operators may prefer command string operations, tailored screen options, and options for operator customization. It is best if operators can get a hands-on trial evaluation of the console capabilities prior to final selection.

*Mark as Analyzed* – Operators will typically be faced with large numbers of alarms that have to be analyzed and cleared. A capability that is usually critical is the ability to selectively keep track of alarms that have been reviewed is very useful.

*Drill Down* – Many network IDS consoles display a high level characterization of events in order to display the large number of alarms that are detected. Operators will usually have to access additional details about each alarm to be able to characterize it properly. It is very desirable for the console to be able to provide the additional levels of information when requested by the operator. As with the operator interface, the types of information desired will typically depend on the skills of the operators.

*Correlation* – In the same vein as drill down features, operators will require tools for correlating (e.g., based on source, destination, type alarms and events) in order to identify and properly characterize intrusions and attacks. This is particularly the case in situations where the incidents are distributed in time or location. The ability of the console to integrate the reporting of various network-based and host-based IDSs and other relevant events is a strong plus, if the operators will use the additional information. Again, as with the operator interface, the types of tools desired will typically depend on the skills of the operators.

*Reporting* – The type of reporting options will depend predominantly on the type of information operators will want to perform their characterization, and the organization's need for reporting to higher levels (e.g., periodic summary reports). It is always desirable to select a console that is capable of generating reports that can be created and disseminated with little extra effort beyond the hour-to-hour and minute-to-minute responsibilities that the operators will have otherwise.

## 6.4.1.5 Considerations for Deployment

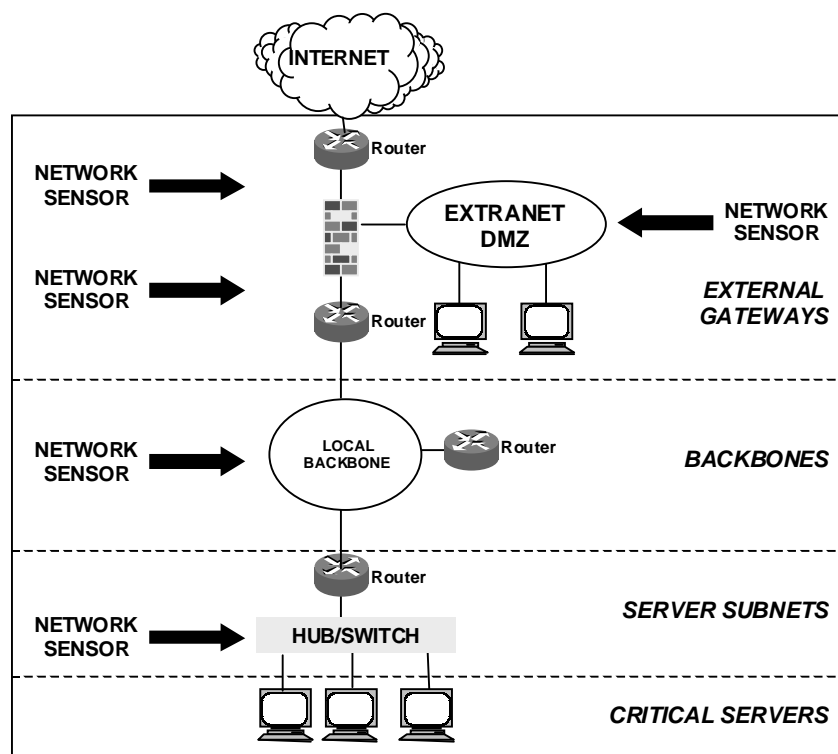
Network architectures present another major challenge for a network IDS. Network switches, which segregate network traffic into specific individual “subnets” reduce network loads across an organization by implementing a form of “need to know” policy among connected computers. Network switches only allow traffic to enter a subnet if it is meant for a computer within that subnet; similarly, they only allow packets out of a subnet that are destined for a computer outside its particular realm.

A network IDS can see only traffic available on the segments on which it is installed. As long as the network IDS is placed on critical segments, it will be able to measure the effectiveness of the security protection mechanisms for the most critical systems and applications. Within an enclave environment, there are a number of possible locations to consider in deploying a network IDS, as depicted in Figure 6.4-2. The challenge is to identify where the traffic of most interest (i.e., highest probability to be used as attack channels) can be monitored.

The external gateways are an obvious candidate in that they allow the IDS to see all of the traffic destined for the enclave. If they are placed outside the firewall, they have access to the raw wide area network traffic (e.g., Internet) without the benefit of filtering by the firewall. If network encryption is used on that traffic, this will offer little if any value. Placing the IDS inside the firewall resolves network encryption issues but will not give any indication of the effectiveness of the firewall operation. Placing sensors at both points and correlating those that are detected outside but blocked by the firewall could provide this additional perspective. Note that either of these locations provides monitoring for external traffic that is destined for the enclave, or internal traffic that is destined for the wide area network. It does not monitor traffic that is only just internal to the enclave.

If an extranet (or what may be referred to as a demilitarized zone or DMZ) is deployed, an IDS on that segment of the network could offer

monitoring of traffic from outsiders to assets structured for an isolated segment of the enclave. The network backbone represents another deployment option. This does provide access to internal traffic on the backbone. However, at this point in the network, consideration should be



**Figure 6.4-2 Network IDS Deployment Options**

given to the traffic speeds and switching technologies employed on those backbones. In some cases (e.g., ATM, Fiber Distributed-Data Interface (FDDI)) the switching technologies and transmission speeds make currently available IDS technologies impractical.

A final placement option is to locate the network IDS on server subnets. This is typically a good option if hubs are used, such that all traffic on the subnet is available at each hub port. If switches are used rather than hubs, this is still a good option if there is a spanning port is available (that allows access to all traffic). If not, the IDS will not have access to all the traffic through the switch and will be ineffective unless deployed between a host; and a switch (or “onto” a host - see below).

There is always a trade off between the possible deployment locations and the number of IDSs to be deployed. Factors include the workload of the operators needed to analyze and characterize the alarms that each IDS identifies, the complexity of correlating the alarms that multiple monitors will detect from the same event, and the costs associated with purchase, installation, operation, and maintenance of the various deployment options.

### 6.4.1.6 Considerations for Operation

As discussed above, most IDS technologies provide the capability to tune the sensor to improve its performance for specific deployments. When an IDS is first deployed, it is prudent to operate the technology for some period depending on the complexity of the deployment to complete this tuning. This provides a means for determining that the IDS is capable of detecting alarms, and that the IDS is installed on the network as intended (by verifying network addresses that are monitored and the direction of traffic).

Tuning enables the IDS to preclude the detection of authorized traffic patterns that may otherwise cause false positive alarm indications. There are two fundamental approaches for tuning. The first approach is to have knowledge a priori of the traffic sources that could trigger false alarms. This could include the addresses of servers (that expect significant traffic), network management station locations (that normally sweep the network), and computers that are remotely located. The IDS can then be configured (tuned) to preclude these from causing an alarm.

While it is desirable to have the information ahead of time, it is often not available. The other approach is to run the IDS and have it find alarms. As alarms are detected, an analyst determines if indeed it reflected an intrusion or a false positive based on normal operation. This form of “discovery” also provides operators an opportunity to become familiar with the technology before it goes on-line operationally.

Tuning should not be thought of as strictly an installation process. This process should be done on a regular basis, to refine and focus the detection mechanisms on real intrusions, and reduce false positives throughout its operation.

## 6.4.2 Malicious Code (or Virus) Detectors

Malicious code can attack authorized local area network (LAN) users, administrators, and individual workstation/personal computer users in numerous ways such as modifying data in transit, replaying (Insertion of Data), exploiting data execution, inserting and exploiting

malicious code, exploiting protocols or infrastructures bugs, and modifying malicious software during production and/or distribution.

Over the past decade, malicious code (also commonly referred to as computer viruses<sup>2</sup>) have gone from an academic curiosity to a persistent, worldwide problem. Viruses can be written for and spread on virtually any computing platform. Typically, viruses are written to effect client personal computers. However, if the personal computer is connected to other machines on a LAN, it is possible for the virus to invade these machines as well. Please refer to Section 6.6 (Malicious Code Protection), for detailed descriptions of the various types of malicious code, potential malicious code attacks and countermeasures, and requirements for malicious code detection products and technologies.

### 6.4.2.1 Technology Overview

Malicious code scanning technologies prevent and/or remove most types of malicious code. The use of malicious code scanning products with current virus definitions is crucial in preventing and/or detecting attacks by all types of malicious code.

There are several basic categories of anti-virus (AV) technologies:

- *Pre-Infection Prevention Products*—a first level of defense against malicious code, used before a system has been attacked
- *Infection Prevention Products*—used to stop replication processes and prevent malicious code from initially infecting the system.
- *Short-Term Infection Detection Products*—used to detect an infection very soon after the infection has occurred
- *Long Term Infection Detection Products*—used to identify specific malicious code on a system that has already been infected for some time, usually removing the malicious code and returning the system to its prior functionality.

Please refer to Section 6.6.5.2, Viruses and E-mail, for a more detailed description of the types of malicious code detection technologies.

### 6.4.2.2 Important Features

When considering the selection of AV technologies, there are a number of features that should be considered. This section identifies important features for selection. The section that follows discussed rationale for the selection of these features. Additional factors to consider when selecting a malicious code detection product can be found in Section 6.6.6 (Selection Criteria).

#### ***Detection Capabilities***

- Data integrity checks.
- Perimeter level scanning for e-mail and Web traffic.
- Does tool exploit malicious mobile code?

---

<sup>2</sup> Throughout the remainder of this section, the general term *virus* will be used to encompass the broader class of malicious code and delivery mechanisms.



- Real-time virus scanning.
- On-demand virus scanning.
- Network packet monitoring.
- Different strains of polymorphic viruses.
- Viruses residing in encrypted messages, compressed files.
- Viruses in different languages (e.g., JAVA, ActiveX, and Visual Basic).
- Trojan horses and worms.

***Updates***

- Can tool upgrade an existing version?
- Are regular updates available?
- Frequency of update releases.
- Response Mechanisms.
- Quarantine at the server level.
- Quarantine at the console level.
- Supply network-based responders.
- Send alerts to network or system administrators.
- Send alerts (in the case of e-mail borne viruses) to sender and receiver(s).

***Platform Considerations***

- What platforms the tool runs on?
- Does tool allow cross platform support?

### 6.4.2.3 Rationale for Selecting Features

When selecting AV products, two important guidelines must be followed. The “best” product may not be good enough by itself. Also, since data security products operate in different ways, one product may be more useful than another in different situations. The following is rationale for evaluating the features of specific technology offerings. Rating each of these categories will allow an organization to choose the best malicious code product for their needs.

### 6.4.2.4 Detection Capabilities

As discussed in Section 6.6.5.2 (Viruses and E-mail), most computer-virus scanners use pattern-matching algorithms that can scan for many different signatures at the same time. Malicious code detection technologies have to include scanning capabilities that detect known and unknown worms and Trojan horses. Most AV products search hard disks for viruses, detect and remove any that are found, and include an auto-update feature that enables the program to download profiles of new viruses so that it will have the profiles necessary for scanning. The viral signatures these programs recognize are quite short: typically, 16 to 30 bytes out of the several thousand that make up a complete virus. It is more efficient to recognize a small

fragment than to verify the presence of an entire virus, and a single signature may be common to many different viruses.

## Updates

Maintaining an effective defense against virus and hostile code threats involves far more than the ability to produce perfect detection rates at a given point in time. With an average of nearly 300 new viruses discovered each month, the actual detection rate of AV software can decline rapidly if not kept current. This AV protection should be updated regularly. As new viruses are discovered, corresponding cures are developed to update protections. These updates should not be ignored. AV systems should do these updates automatically, reliably, and through a centrally controlled management framework. To stay current, these scanning programs must be updated when new viral strains are found and anti-viral codes are written. Most computer-virus scanners use pattern-matching algorithms that can scan for many different signatures at the same time. This is why enterprise-class AV solutions must be able to offer timely and efficient upgrades and updates across all client and server platforms.

Often, in large enterprise environments, a typical acquisition and deployment strategy is to deploy one brand of AV software at end-user workstations and a different vendor's product on the e-mail, file, and application server environments. This broadens the spectrum of coverage because in any given instance, one vendor is typically ahead of another in releasing the latest round of virus signature discoveries.

## Response Mechanisms

Once malicious code has been detected, it must be removed. One technique is simply to erase the infected program, but this is a harsh method of elimination. Most AV programs attempt to repair infected files rather than destroy them. If a virus-specific scanning program detects an infected file, it can usually follow a detailed prescription, supplied by its programmers, for deleting viral code and reassembling a working copy of the original. There are also generic techniques that work well for known and unknown viruses. One method is to gather a mathematical fingerprint for each program on the system. If a program subsequently becomes infected, this method can reconstitute a copy of the original. Most tools perform scanning for viruses, but all do not detect and remove Trojan Horses, worms, and malicious mobile code upon all levels of entry. Most currently available AV tools do not have the same capabilities when responding across a network. Additional countermeasures related to malicious code can be found in Section 6.6.4 (Potential Countermeasures).

## Platform Considerations

The computers to run this software must meet the hardware and software requirements specified by the manufacturer. The malicious code protection software should function properly and perform its duties without failing or interfering with other applications running on the same system.

## 6.4.2.5 Considerations for Deployment

Defense in depth dictates that any virus protection must be implemented across the enterprise. This means installing and managing AV software on every system. Some advocate only installing AV only on edge devices such as servers, firewalls, and gateways. Defense against viruses is only as good as its weakest link and if one system can be compromised, then the entire enterprise is at risk.

Centralized management for the AV capabilities with a common set of policies is strongly recommended. Though some vendor offerings cater to end-users being held responsible for security mandates, this can lead to more and varied security holes. What most often happens is that end users have their session interrupted with a pop-up screen that tells them their files are about to be scanned or that they are about to receive an AV update. Many users tend to override the update manually, as it is distracting.

## 6.4.2.6 Considerations for Operation

Most AV technologies provide a means for sending responses or alerts at the server level, and some at the console level. It is always desirable to notify anyone that may have been infected that malicious code has been detected. This should include system and network administrators. If it is encountered in e-mail transactions, it is desirable to notify the sender and recipient. If it is found on a file system that knows the file owner, he should be notified. In general, anyone that could be notified should be.

## 6.4.3 Discussion of Typical Bundling of Capabilities

At one point, network monitors were offered as stand-alone devices. Vendors may prefer to offer these technologies as appliances, being sold with what is otherwise a COTS computer system, at an inflated price. There are a number of offerings combine these monitors with firewalls, routers, vulnerability scanners and the like as a means for vendors to leverage existing market positions to gain market share in related areas. Another trend that is becoming popular is for larger vendors to offer integrated architecture approaches, where they combine a number of related technologies as a bundled offering. They tend to prefer custom rather than standard interfaces to preclude the merging of other vendor offerings. This offers so-called “complete solutions”, however, it tends to lock the buyer into one particular product suite. While this often sounds attractive, it is often valuable to be able to integrate various technologies together in order to take advantage of the detection capabilities available from the different implementations.

There is a natural linkage of these monitoring technologies with Enterprise Security Management (ESM) systems, and vendors have been talking about the integration for some time. However, there has been, and there is little evidence to suggest that this integration will be realized in the foreseeable future.

## 6.4.4 Beyond Technology Solutions

While the focus of the IATF is on technology solutions, there are important operational aspects of effective network monitoring are also critical to an effective IA solution. The framework recommends the following guidance:

### ***Operational Planning***

- Develop intrusion detection and anti-viral-related requirements as an integral part of the enterprise security policy.
- Assess the ability of system administration personnel to perform intrusion detection and related vulnerability scanning.
- Consult with experienced intrusion detection and vulnerability scanning personnel regarding the best approach.
- Consult with the appropriate legal council regarding affected personnel rights and procedures, as discussed below.
- Provide for adequate technical and legal training of all involved personnel.
- Acquire software and expertise from a high integrity vendor.
- Perform network monitoring consistent with the enterprise security policy.
- Tightly couple vulnerability scanning and intrusion detection activities.

### ***Intrusion Detection Activities***

- Look for intrusion evidence based on found vulnerabilities; use intrusion evidence to find and correct vulnerabilities.
- Provide and monitor bogus sites/services/information. Possibly monitor intrusions through known vulnerabilities to satisfy prosecution requirements in conjunction with the appropriate legal authorities.
- Perform intrusion responses that are approved by the appropriate authority.

### ***Network Malicious Code Detection Activities***

- Select and deploy virus scanning capabilities that are consistent with the location, functions, and capabilities.
- Acquire or download the appropriate anti-viral software from a high integrity source, and acquire any necessary hardware (such as an ancillary firewall dedicated to virus scanning incoming or outgoing traffic).
- Institute enterprise-wide anti-viral training and procedures.
- Scan consistently based on time and/or events.
- Follow up on all indications of potential contamination (as defined in the security policy and anti-viral procedures for the enterprise).
- Update anti-viral software and hardware as appropriate, (e.g., consistent with new releases of anti-viral software and specific experiences throughout the enterprise).

### ***General Activities***

- Archive (within any legal constraints) audit and intrusion information, and correlate with vulnerability scan information.
- Keep authorities apprised of all activities, ensuring that any legal rights are not violated.
- Continuously repeat steps, as appropriate.

### ***Privacy Concerns***

Organizations may own the intellectual property of employees and may also legally restrict computer activities to only those approved by management. A common practice is to present this warning to all computer users as part of the normal login message. This does not mean that ALL managers in an enterprise own ALL of the transactions of ALL of the employees. Especially unclear is how to handle the conflict that arises between privacy and monitoring. Use of IDSs and system monitoring tools requires caution. Sniffers that search for key words in messages (e.g., “attack,” “weakness,” or “confidentiality”) as a standard set of *watchwords* may find them used in an appropriate manner depending on the type of correspondence. Audit trail reports may contain full command strings (including parameters). Knowing that an employee is sending several messages to a particular department (e.g., Human Resources) may be an infringement to his privacy. It is important to refer privacy concerns to the appropriate legal and policy organizations for the enterprise prior to deployment and use of these technologies.

## **6.4.5 For More Information**

The source materials used in the preparation of this section provide an excellent base of knowledge of relevant technologies from which to draw. A number of additional sources of information exist. This section of the framework focuses on on-line sources since they tend to offer up-to-date information. These include the following:

### **6.4.5.1 IATF Executive Summaries**

An important segment of the Information Assurance Technical Framework (IATF) is a series of “Executive Summaries” that are intended to provide summary implementation guidance for specific case situations. These offer important perspectives on the application of specific technologies to realistic operational environments. While these are still being formulated, they will be posted on the IATF web site <http://www.iatf.net> as they become available. [1]

### **6.4.5.2 Protection Profiles**

The National Security Telecommunications and Information Systems Security Policy (NSTISSP) No.11 provides the national policy that governs the acquisition of IA and IA-enabled information technology products for national security telecommunications and information systems. This policy mandates that effective January 2001 preference be given to products that are in compliance with one of the following:

- International Common Criteria for Information Security Technology Evaluation Mutual Recognition Arrangement

## UNCLASSIFIED

Network Monitoring Within Enclave Boundaries and External Connections  
IATF Release 3.0—September 2000

- National Security Agency (NSA)/National Institute of Standards and Technology (NIST) National Information Assurance Partnership (NIAP)
- NIST Federal Information Processing Standard (FIPS) validation program.

After January 2002, this requirement is mandated. Department of Defense (DoD) Chief Information Officer (CIO) Guidance and Policy Memorandum No. 6-8510, Guidance and Policy for Department of Defense Global Information Grid Information Assurance references this same NSTISSP No. 11 as an acquisition policy for the Department.

- The International Common Criteria and NIAP initiatives base product evaluations against Common Criteria Protection Profiles.
- NSA and NIST are working to develop a comprehensive set of protection profiles for use by these initiatives. An overview of these initiatives, copies of the Protection Profiles, and status of various products that have been evaluated are available at the NIST web site <http://niap.nist.gov/>. [2]

### 6.4.5.3 Independent Third Part Reviewers of Relevant Vendor Technologies

ICSA Net Security Page [www.icsa.net](http://www.icsa.net)

Talisker's Intrusion Detection Systems [www.networkintrusion.co.uk/](http://www.networkintrusion.co.uk/)

Network Computing – The Technology Solution Center  
[www.nwc.com/1023/1023f12.html](http://www.nwc.com/1023/1023f12.html)

Paper on CMDS Enterprise 4.02 <http://www.intrusion.com/Products/enterprise.shtml> (ODS Networks has changed its name to Intrusion.com)

PC Week On-Line [www.zdnet.com/pcweek/reviews/0810/10sec.html](http://www.zdnet.com/pcweek/reviews/0810/10sec.html)

### 6.4.5.4 Overview of Relevant Research Activities

Coast Homepage – Perdue University [www.cs.purdue.edu/coast](http://www.cs.purdue.edu/coast)

UC Davis <http://www.gidos.org/>

UC Davis <http://seclab.cs.ucdavis.edu/>

### 6.4.5.5 Overview of Selected Network Monitor Vendor Technologies

Axent Technologies [www.axent.com](http://www.axent.com)

cai.net <http://www.cai.net/>

Cisco Connection Online [www.cisco.com](http://www.cisco.com)

CyberSafe Corporation [www.cybersafe.com](http://www.cybersafe.com)

Internet Security Systems [www.iss.net](http://www.iss.net)

**UNCLASSIFIED**

Network Monitoring Within Enclave Boundaries and External Connections  
IATF Release 3.0—September 2000

Network ICE [www.networkice.com](http://www.networkice.com)

# References

1. Information Assurance Technical Framework (IATF) <http://www.iatf.net> .
2. National Institute of Standards and Technology <http://niap.nist.gov/>.

# Additional References

- a. Amoroso, Edward, Intrusion Detection. Intrusion.Net Books. 1999.
- b. Escamilla, Terry. Intrusion Detection, Network Security Beyond the Firewall. Wiley Computer publishing. 1998.
- c. Northcutt, Stephen. Network Intrusion Detection, An Analyst's Handbook. New Riders Publishing. 1999.
- d. Snapp, Steven R., et al. A System for Distributed intrusion Detection. IEEE CH2961-1/91/0000/0170. 1999
- e. Balasubramaniyan, J. S., et al. An Architecture for Intrusion Detection using Autonomous Agents. COAST Technical Report. 11 June 1998.
- f. AXENT Technologies, Inc. Intruder Alert 3.5 IDS Review Guide, May 2000.
- g. AXENT Technologies, Inc. Everything You Need to Know About Intrusion Detection, 1999.
- h. Schneider, Sondra, et al. Life After IDS. Information Security Magazine. Volume 2, Number 9. September 1999.
- i. Graham, Robert. New Security Trends for Open Networks. SC Magazine. October 1999.
- j. SC Magazine. Intrusion Detection. June 2000.
- k. Information Assurance Technology Analysis Center (IATAC). Tools Report on Intrusion Detection. Defense Technical Information Center. December 1999.
- l. Maes, V. How I Chose an IDS. Information Security Magazine. Volume 2, Number 9. September 1999.
- m. 13. Concurrent Technologies Corporation. Attack Sensing, Warning, and Response (ASW&R) Trade Study Report Intrusion Detection System. Report No. 0017-UU-TE-000621. April 14, 2000.
- n. Information Assurance Technology Analysis Center (IATAC). Tools Report on Vulnerability Analysis Information. Defense Technical Information Center. March 15, 2000.
- o. Ulsch, Macdonnell and Joseph Judge. Bitter-Suite Security. Information Security Magazine. Volume 2, Number 1. January 1999.
- p. Concurrent Technologies Corporation. Attack Sensing, Warning, and Response (ASW&R) Baseline Tool Assessment Task Anti-Virus Trade Study Report. Report No. 0017-UU-TE-000623. April 13, 2000.



**UNCLASSIFIED**

Network Monitoring Within Enclave Boundaries and External Connections  
IATF Release 3.0—September 2000

- q. Department of Defense (DoD) Chief Information Officer (CIO) Guidance and Policy Memorandum No. 6-8510, Guidance and Policy for Department of Defense Global Information Grid Information Assurance
- r. National Security Telecommunications and Information Systems Security Policy (NSTISSP) No. 11. National Policy Governing the Acquisition of Information Assurance (IA) and IA-Enabled Information Technology (IT) Products. January 2000.

**UNCLASSIFIED**

Network Monitoring Within Enclave Boundaries and External Connections  
IATF Release 3.0—September 2000

**This page intentionally left blank.**

**UNCLASSIFIED**

## 6.5 Network Scanners Within Enclave Boundaries

As discussed in Section 6.4, Network Monitoring Within Enclave Boundaries and External Connections, on-line network monitoring technologies provide a critical layer of defense within enclave boundary protection. In addition to the network monitoring technologies, is another class of technologies, referred to as network scanners, it can also be deployed to improve overall security posture. The Framework makes the distinction between these scanners and network monitoring devices. Monitors typically operate in near real time and are based on network traffic (or related characteristics) for their focus. Monitors tend to measure the effectiveness of the network's protection services because they are subject to attempted exploitation. This is somewhat of an "after the fact" measure, not a preventative measure. Scanners, on the other hand, are preventative measures. Typically, they operate periodically (or on demand) and examine systems for vulnerabilities that an adversary could exploit, measuring effectiveness of the infrastructure's system's infrastructure protection.

The local environment is the logical place for addressing these network assessment technologies. Scanning can be performed at the network boundary or host levels. This segment of the Information Assurance Technology Framework (IATF) specifically considers Network Vulnerability Scanner and War Dialer technologies that are germane to the enclave environment. Please refer to Section 7.2, Host-based Detect and Respond Capabilities within Computing Environments, of the Framework for guidance on the use of similar technologies that are more suitable for deployment at the host level.

Unlike the near real-time network monitoring technologies addressed in Section 6.4, Network Monitoring Within Enclave Boundaries and External Connections, network assessment technologies are typically executed in a periodic or on-demand manner, providing perspectives on the posture of a local environment. Section 8.2, Detect and Respond as a Supporting Element, of the Framework provides a perspective on an overall detect and response infrastructure; however, because these assessments typically focus on the local level, they tend not to interact with or be particularly relevant to a broader network infrastructure.

### 6.5.1 Network Vulnerability Scanners

Periodic or on-demand network assessment tools are adept at finding security holes at boundary points devices or on network hosts within an enclave environment, hopefully before an attacker does. They accomplish this effort by discovering known vulnerabilities in host or network system components and improper configurations visible from the network that create the potential for unauthorized access or exploitation or are counter to enterprise policies.

#### 6.5.1.1 Technology Overview

Vulnerability analysis tools help automate the identification of vulnerabilities in a network or system. Network-based vulnerability scanners take inventory of all devices and components

within the network infrastructure. These scanners operate over a network “against” target nodes by probing and examining the network components and hosts in order to identify vulnerabilities that are typically “visible” to their network connection. They look to identify network services that allow uncontrolled access, contain buffer control vulnerabilities, violate possible trust privileges that allow potential exploitation, and contain weaknesses in network component (e.g., router, firewall, and Web server) configurations.

A scanner probes for weaknesses by comparing data about a network configuration against its database of known vulnerabilities. These components, the network configuration, and the various versions of the software controlling the network are examined and compared with a database of known vulnerabilities. Network vulnerability scanners fall within one or more of the following classes.

## Simple Vulnerability Identification and Analysis

A number of tools have been developed that perform relatively limited security checks. These tools may automate the process of scanning Transmission Control Protocol/Internet Protocol (TCP/IP) ports on target hosts, attempting to connect to ports running services with well known vulnerabilities and recording the response. They also may perform secure configuration checks for specific system features. The user interface of these tools is likely to be command-line based, and the reporting may include limited analysis and recommendations. The tools are likely to be “freeware.”

## Comprehensive Vulnerability Identification and Analysis

More sophisticated vulnerability and analysis tools have been developed that are fairly comprehensive in terms of the scope of vulnerabilities addressed, the degree of analysis performed, and the extent of recommendations made to mitigate potential security risks. Many of these tools also provide a user-friendly graphical user interface, or GUI.

## Password Crackers

Password cracker tools attempt to match encrypted forms of a dictionary list of possible passwords with encrypted passwords in a password file. This is possible because the algorithm used to encrypt operating system’s passwords is public knowledge. An attacker or insider would run these tools after successfully gaining access to the system in order to acquire a higher privileged level such as root. These tools allow operators to verify compliance with password selection policies. Many tools from the previous category have integrated password-cracking modules.

## Risk Analysis Tools

Risk Analysis tools typically provide a framework for conducting a risk analysis but do not actually automate the vulnerability identification process. These tools may include large databases of potential threats and vulnerabilities along with a mechanism to determine, based on user input (typically query/response scripts), cost-effective solutions to mitigate risks. The vulnerabilities identified using a vulnerability analysis tool may be input into a risk analysis tool to assist in determining the overall risk to the system. Or vice versa, vulnerabilities

predicted by a risk analysis tool can be specifically targeted for confirmation using vulnerability scanning tools.

### 6.5.1.2 General Considerations for Use

Network vulnerability scanners operate across the network to identify weaknesses in a connected system's security scheme, exploitation of which would negatively effect the confidentiality, integrity, or availability of the system or its information. These scanners are easy to install and can run a wide variety of attacks on a network to determine the network's resilience to each attack. However, a scanner takes a snapshot of the network and does not operate in real time, often requiring post-capture analysis to understand and implement any mitigation approaches that may be required. Typically, Local Area Network (LAN) administrators do not use them on a day-to-day basis.

Scanners work by either examining attributes of objects or by emulating an attacker. To act as a hacker, a scanner can execute a variety of attack scripts. Because these can look (and act) like real attacks, it is important to consider what and when scans are performed. Otherwise, it is possible that the scanner could cause as much impact to the network as an actual incident. Coordination with network operations staff is critical, particularly in environments that implement real-time intrusion detection techniques. However, another use of such scanners is a "live" test and readiness evaluation of intrusion detection and incident response processes and procedures for an enterprise environment.

The vulnerability scanners will detect only objects it is configured to scan. If the scanner is not configured and set up properly, vulnerabilities may exist that will not be identified. This may be of less value than not performing the scans because it may offer a false sense of the adequacy of the network's resiliency to attacks.

### 6.5.1.3 Important Features

When considering the selection of a network-based vulnerability scanners, a number of features should be considered. This section identifies important features for selection. The section that follows discussed rationale for the selection of these features.

#### Scanning Capabilities

- Does the tool offer an ability to add custom scanning routines to look for site- or technology-specific weaknesses of concern.

#### Response Mechanisms

- Automatic shutoff of vulnerable ports of entry.

#### User Interfaces

- Does the tool have a GUI for number entry, dialing status, and call results.
- Can reports be viewed in real time.

## Reporting Capabilities

- Does it offer automatic alerting when new non-network ports are detected.
- Are all system answers logged in a database or file.
- Is there an updated database of network numbers with which to compare newly identified numbers.
- Does the database automatically combine logged information and place it in a report format.
- Does the tool provide suggested mitigation approaches for discovered vulnerabilities.

## Platform Compatibility

- What are the platforms (operating systems) on which the tool will run.
- Does it use executables.
- Does it support scripts or macros.

### 6.5.1.4 Rationale for Selecting Features

The type and level of detail of information provided among tools varies greatly. Although some can identify only a minimal set of vulnerabilities, others can perform a greater degree of analysis and provide detailed recommended mitigation approaches. Scanner technologies should be selected that cover the gamut of vulnerabilities for the given environment and system platforms. Technologies that offer a comprehensive library of vulnerabilities periodically updated by the vendor should be selected. Capabilities including grouping of nodes into scan groups and customized scan options may be valuable for larger environments.

Some scanner technologies offer features that are useful depending on the training and skill levels of operators that will be using them. Depending on the planned usage of the scanner and the skills of the operators available, it is often desirable to select technologies that can be tuned to ignore some false positives. It is also desirable to select features that enable the scanner to be tuned for important application environments, such as database environments, Web Server environments, file server environments, firewalls, etc., since such profiles may differ based on the functions provided.

#### Source

- Has the tool been developed by the Government (or under Government sponsorship); if so, is it reserved; can your organization obtain authorization for its use.
- Is the tool available from a reputable vendor.
- Is the tool in the public domain (e.g., freeware from the Internet); if so, is source code available.

## Scanning Capabilities

The type and level of detail of information provided among tools varies greatly. Although some can identify only a minimal set of vulnerabilities, others can perform a greater degree of analysis and provide detailed recommended mitigation approaches.

## Response Mechanisms

Assessment tools will continue to evolve in usability, with some vendors offering click-and-fix solutions. The assessment software flags vulnerabilities in terms of the risk posed to the network and the ease of the fix. Some technologies can generate trouble tickets to trigger a manual response. They may offer an ability to change policies in firewalls and other enclave boundary defense mechanisms. Some identify patches that should be installed. Some offer to obtain and install patches. Although installing patches is feasible, allowing the security administrator the ability to undertake these tasks and the difficulty of undoing configuration changes should leave customers wary of this function. Such features should be considered in light of an environment's existing configuration management policies and procedures.

## User Interfaces

Most scanners enable the operator to configure what network elements are to be scanned, and when the scans are to occur. Typically, scanners are preconfigured with lists of vulnerabilities and can operate without customization. Some technologies allow operators to customize the vulnerabilities the scanner will investigate. Usually the results are sorted into a file that can be accessed upon demand to review the results. More recently developed tools provide user-friendly front ends and sophisticated reporting capabilities.

## Reporting Capabilities

Old products inundated customers with phonebook-size reports on all the various vulnerabilities that the network faced. New products have database interfaces that prioritize vulnerabilities and allow network managers to deal with the network's problems in a logical manner. Many generate reports that are Web-enabled with hot-links and other "labor savers."

## Platform Compatibility

The computers to run this software must meet the hardware and software requirements specified by the manufacturer. The vulnerability scanner software should function properly and perform its duties without failing.

## 6.5.2 War Dialers

Firewalls and other enclave boundary protection devices can create a level of defense against network attacks that adversaries have to defeat. However, as the trend continues toward borderless networks, machines with attached modems are often scattered throughout organizations. When modems are installed on telephone lines connected to the data network, firewalls are no longer the only access port to the network, and thus cannot detect or control ALL of the data traffic that is traveling in or out of the network. The result is that "back doors"

are created that offer alternative, unprotected portals for adversaries to exploit, as depicted in Figure 6.5-1. Analysts estimate that the bulk of damaging hacks on corporate networks comes over modem connections that are not secure. One technology called War Dialers, is a specific form of network vulnerability scanners.

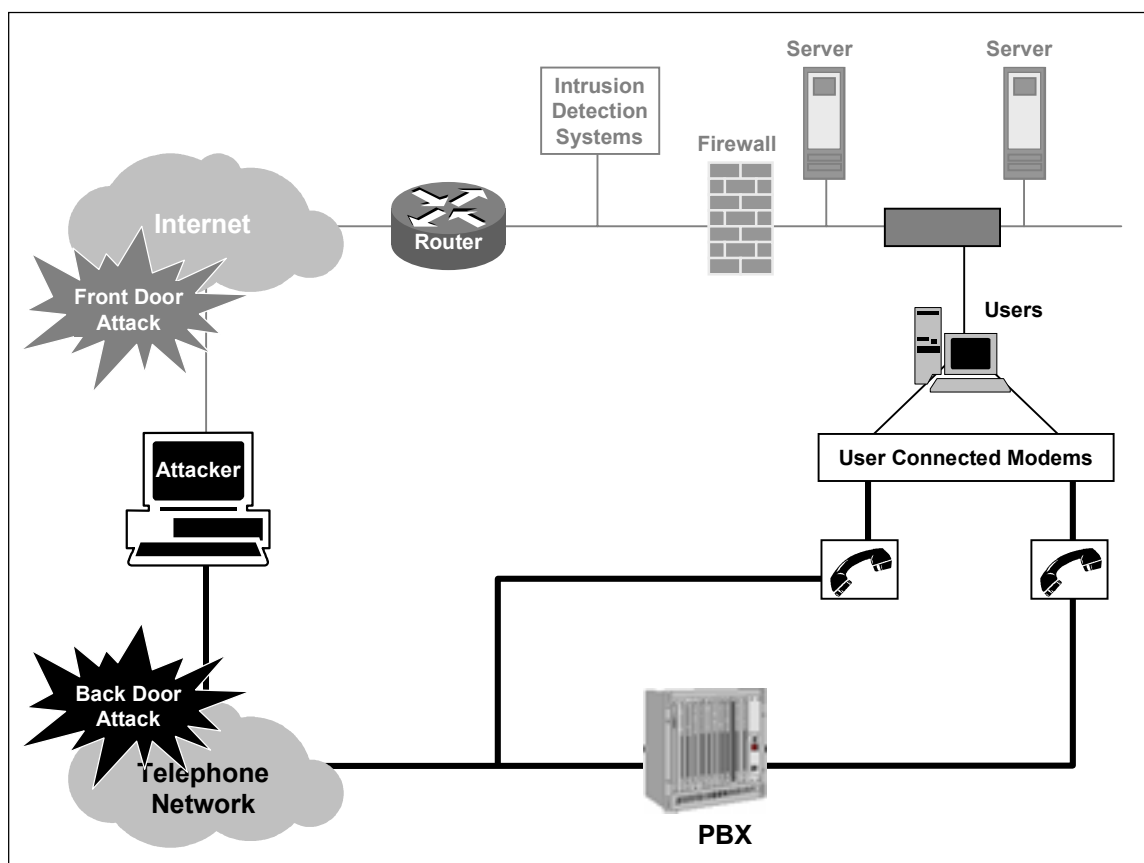


Figure 6.5-1 Back door Attacks Through Telephone Networks

### 6.5.2.1 Technology Overview

Most commonly, War Dialers are associated with hackers. Most hackers target organizations because they rarely control the dial-in ports as strictly as a firewall. One way of combating intrusions accomplished by hackers is to use the same type of scanning tool as a defensive mechanism.

A War Dialer consists of software that dials a specific range of telephone numbers looking for modems that provide a login prompt. The tools, at a minimum, record the modem numbers and login screen, but can also be configured to attempt brute force, dictionary-based, login attempts. Visibility into telephone networks is provided by identifying modem, fax, or voice tones and characterizing security behaviors. This process allows identification of network vulnerabilities.

War Dialers call a given list or range of telephone numbers and record those that answer with handshake tones. Those handshake tones may be characterized as entry points to computer or



telecommunications systems. Some of these programs have become quite sophisticated, and can now detect modem, fax, or Private Branch Exchange (PBX) tones and log each one separately. A block of specified numbers is attempted and any modems found in that block are noted.

## 6.5.2.2 General Considerations for Use

Remote access into most organizations' information systems is usually performed through ordinary telephone lines. The lack of visibility into the use of telephone networks makes it possible for any user to connect to an entire private data network via a modem. These telephone lines must be thought of as ports of entry for possible network attacks and intrusions. When an enclave does not deploy protection mechanisms that effectively secure or monitor telephone networks, intruders can gain access to proprietary information; existing security systems remain blind to unauthorized activity. War Dialers are an effective way to identify unsecured modems. Along with a strong modem policy describing the need for modem registration and PBX controls, War Dialer scanning can help an organization defend itself against such dangers. Use of this type of technology can help an enterprise to identify those vulnerable backdoors before an attack occurs. Once identified, those backdoors can be closed or some type of security plan created to preclude use of that particular point of entry.

## 6.5.2.3 Important Features

When considering the selection of a War Dialer technology, a number of features should be considered. This section identifies important features for selection. The section that follows discusses rationale for the selection of these features.

### Scanning Capabilities

- Identification of every dial-up system.
- Facsimile machine detection.
- Multi-modem scanning.
- Brute force username and/or password guessing (code cracking).
- Support terminal emulation to allow tool to enable access to mainframe computers.
- Built-in knowledge of various dial-in authentication technologies.

### Response Mechanisms

- Automatic shutoff of vulnerable ports of entry (interface to telephone network).

### User Interfaces

- Does the tool have a GUI for number entry, dialing status, and call results.
- Can reports be viewed in real time.

## Reporting Capabilities

- Automatic alerting when new non-network ports are detected.
- Are all system answers logged in a database or file.
- Is there an updated database of network numbers with which to compare newly identified numbers.
- Does the database automatically combine logged information and place it in a report format.

## Platform Compatibility

- What platforms (operating systems) will the tool run on.
- Does it use executables.
- Does it support scripts or macros.

## Source

- Has the tool been developed by the Government (or under Government sponsorship); if so, is it reserved; can your organization obtain authorization for its use.
- Is the tool available from a reputable vendor.
- Is the tool in the public domain (e.g., freeware from the Internet); if so, is source code available.

## 6.5.2.4 Rationale for Selecting Features

War Dialers identify known modems, modem banks, and communication servers; compare discovered modem configuration data against predefined modem configurations; and alert administration when a vulnerable port of entry has been detected. The major discriminator is how well each product performs these functions.

It is often difficult to determine the true nature of the features that are provided in a particular technology offering (beyond strict vendor claims). It is always advisable to seek test results of reputable, independent third-party laboratories. When these are available, they should be an important consideration in any technology selection. A number of organizations provide these types of results.

## Scanning Capabilities

It is important that the War Dialer be capable of uncovering and characterizing all back doors on the network, because each represents a potential unprotected portal for an adversary. Thus, beyond simply identifying when a modem responds to an incoming call on each telephone line specified, it is possible to uncover when computers serving as facsimile machines and modem banks are encountered. Further, the ability to emulate a terminal (to enable access to mainframe computers) and apply password cracking mechanisms provide valuable information regarding how susceptible the identified parts actually are, supporting efforts to prioritize those that require earlier resolution. The more extensive scanning capabilities a tool offers will enable a more thorough and reliable report on the actual posture of the network.

## Response Mechanisms

For the most part, War Dialers typically report on back doors they have uncovered. However, technologies are available that can automatically shut off vulnerable ports of entry. Care should always be taken when selecting any automated response. In this case, shutting down a remote access port may have negative effects on operational capabilities.

## User Interfaces

Most scanners enable the operator to enter telephone numbers and provide dialing status and call results. Usually the results are stored in a file that can be accessed upon demand to review the results. Depending on the skills of the intended operator, it may be desirable to select a tool that offers a user-friendly interface. Recently developed tools provide a user-friendly user interface for number entry, dialing status, and call results.

## Reporting Capabilities

Again, based on the intended manner in which the War Dialer is operated, it may be desirable to select features that provide automatic alerting when new non-network ports are detected. If reports of the results of War Dialer scans are required by the organization, consideration should be given to technologies that offer the capability for the database to automatically combine logged information and place it in a report format. If the enterprise allows selected remote access ports to remain operational, operators may be concerned primarily with new ports that were not reported previously. In this situation, consideration should be given to technologies which are able to update the database of network numbers with which to compare newly identified numbers.

It is important to ensure that the selected technology logs all system answers in a database or file. If the operator will be monitoring the results of the War Dialer assessment during its operation, it will be important to consider technologies where reports can be viewed in real time.

## Platform Compatibility

The computers to run this software must meet the hardware and software requirements specified by the manufacturer. The malicious code protection software should function properly and perform its duties without failing.

## Source

A number of War Dialers have been developed by the Government (or under Government sponsorship). If one of these is selected, it may be reserved for use only by selected communities. In these situations, it is necessary to determine if your organization can obtain authorization for its use.

A wide array of War Dialers is available as freeware or shareware. These are regarded as hacker tools and are an open source via the Internet. Many commercial scanners dial only predetermined numbers in a telemarketing atmosphere. Commercial products are preferred because they tend to offer technical support mechanism; typically, no reliable means exist for support for freeware and/or shareware. Overall, the functions are the same, but technical

support, better reporting styles, and more attractive GUIs can be found with the commercial products offered today.

Care should be taken when using any software obtained from the public domain (e.g., freeware from the Internet). The software should be scanned carefully for potential malicious code. If source code is not available, its use is *NOT* recommended.

### 6.5.3 Considerations for Deployment

The same considerations for placement of network monitors, as discussed in Section 6.4 (Network Monitoring Within Enclave Boundaries and External Connections), are in general applicable for deploying network scanners. Network switches, which segregate network traffic into specific individual “subnets,” reduce network loads across an organization by implementing a form of “need-to-know” policy among connected computers. Network switches allow traffic to enter a subnet only if it is meant for a computer within that subnet; similarly, packets are only allowed out of a subnet that are destined for a computer outside its particular realm.

Network scanners only can find vulnerabilities that they can see based on the segments on which they are installed. As long as the network scanner is placed on critical segments, it will be able to measure the effectiveness of the security protection mechanisms for the most critical systems and applications. Within an enclave environment, a number of possible locations should be considered in deploying a network intrusion detection system (IDS). The challenge is to identify the locations where the potential vulnerabilities are of most interest. This is often considered from the view of potential attacker sources that are of concern. For example, if the concern is from a hacker from the Internet, the scanner should be structured to look at the network from that vantage point. If the concern is for insider threats, that vantage point should be considered. Because the scanners can operate on demand, they can be used in one location and then moved to another to determine the overall security posture of a network environment.

### 6.5.4 Considerations for Operation

Assessment frequency is a factor of how often network changes are made and the security policy for the enterprise. Depending on the organization, assessments may take place quarterly, monthly, weekly, or even daily. Some service providers offer scanning services on a subscription basis, ensuring that assessments occur regularly.

### Discussion of Typical Bundling of Capabilities

At one point, network monitors were offered as stand-alone devices. Vendors may prefer to offer these technologies as appliances, being sold with what is otherwise a Commercial-Off-The-Shelf (COTS) computer system, at an inflated price. A number of offerings combine these monitors with firewalls, routers, vulnerability scanners, and the like as a means for vendors to leverage existing market positions to gain market share in related areas. Another trend that is becoming popular is for larger vendors to offer integrated architecture approaches, where they combine a number of related technologies as a bundled offering. Vendors tend to prefer custom rather than standard interfaces to preclude the merging of other vendor offerings. This

offers so-called “complete solutions;” however, it tends to lock the buyer into one particular product suite. Although this often sounds attractive, it is valuable to be able to incorporate various technologies to take advantage of the detection capabilities available from the different implementations.

There is a natural linkage of these monitoring technologies with Enterprise Security Management (ESM) systems, and vendors have been discussing the integration for some time. However, there is little evidence to suggest that this integration will be realized in the foreseeable future.

## 6.5.5 Beyond Technology Solutions

Although the focus of the IATF is on technology solutions, important operational aspects of effective network scanning that are critical to an effective IA solution. Network scanning is the primary means of assessing the security of the network. The functions performed by the scanner should be tailored to the network configuration and environment, together with the applications performed by the protected network. The framework recommends the following guidance for network scanners:

- Develop network scanning requirements as an integral part of the enterprise security policy.
- Scan your network consistent with the guidance listed for intrusion detection and response, using the best available scanners.
- Assess the results in light of your security policy.
- Adjust and counter identified deficiencies relative to your policy. This may include patches, changes in configuration, changes in procedures, or better enforcement of procedures such as the use of good passwords that change frequently.
- Repeat the process regularly.

## 6.5.6 For More Information

The list of reference materials used in preparing this section provides an excellent base of knowledge from which to draw on relevant technologies. A number of additional sources of information exist. This section of the Framework focuses on on-line sources because they tend to offer up-to-date information. These include the following.

### 6.5.6.1 IATF Executive Summaries

An important segment of the IATF is a series of “Executive Summaries” that provides summary implementation guidance for specific case situations. These summaries offer important perspectives on the application of specific technologies to realistic operational environments. Although these are still being formulated, they will be posted on the IATF Web site [www.iatf.net](http://www.iatf.net) as they become available. [1]

## 6.5.6.2 Protection Profiles

The National Security Telecommunications and Information Systems Security Policy (NSTISSP) Number 11 provides the national policy that governs the acquisition of Information Assurance (IA) and IA-enabled information technology products for national security telecommunications and information systems. This policy mandates that effective January 2001, preference be given to products that are in compliance with one of the following.

- International Common Criteria for Information Security Technology Evaluation Mutual Recognition Arrangement.
- National Security Agency/National Institute of Standards and Technology (NSA/NIST) National Information Assurance Partnership (NIAP).
- NIST Federal Information Processing Standard (FIPS) validation program.

After January 2002, this requirement is mandated. Department of Defense (DoD) Chief Information Officer (CIO) Guidance and Policy Memorandum No. 6-8510, Guidance and Policy for Department of Defense Global Information Grid Information Assurance references this same NSTISSP Number 11 as an acquisition policy for the Department.

The International Common Criteria and NIAP initiatives base product evaluations against Common Criteria Protection Profiles. NSA and NIST are developing a comprehensive set of protection profiles for use by these initiatives. An overview of these initiatives, copies of the Protection Profiles, and the status of various products that have been evaluated are available at the NIST Web site <http://niap.nist.gov/>. [2]

## 6.5.6.3 Independent Third Part Reviewers of Relevant Vendor Technologies

ICSA Net Security Page [www.icsa.net](http://www.icsa.net)

Talisker's Intrusion Detection Systems [www.networkintrusion.co.uk/](http://www.networkintrusion.co.uk/)

Network Computing – The Technology Solution Center  
[www.nwc.com/1023/1023f12.html](http://www.nwc.com/1023/1023f12.html)

Paper on CMD5 Enterprise 4.02 [www.ods.com/downloads/docs/Cmd5-us.pdf](http://www.ods.com/downloads/docs/Cmd5-us.pdf) (ODS Networks has changed its name to Intrusion.com)

PC Week On-Line [www.zdnet.com/pcweek/reviews/0810/10sec.html](http://www.zdnet.com/pcweek/reviews/0810/10sec.html)

## 6.5.6.4 Overview of Relevant Research Activities

Coast Homepage – Perdue University [www.cs.purdue.edu/coast](http://www.cs.purdue.edu/coast)

UC Davis [www.seclab.cs.ucdavis.edu/cidf](http://www.seclab.cs.ucdavis.edu/cidf)

UC Davis [www.seclab.cs.ucdavis.edu](http://www.seclab.cs.ucdavis.edu)

## 6.5.6.5 Overview of Selected Network Scanner Vendor Technologies

Axent Technologies [www.axent.com](http://www.axent.com)

cai.net <http://www.cai.net/>

Cisco Connection Online [www.cisco.com](http://www.cisco.com)

CyberSafe Corporation [www.cybersafe.com](http://www.cybersafe.com)

Internet Security Systems [www.iss.net](http://www.iss.net)

Network ICE [www.networkice.com](http://www.networkice.com)

## 6.5.6.6 Overview of Selected War Dialer Technologies

VerTTeX Software [www.verttex.com](http://www.verttex.com)

The Hackers Choice [www.infowar.co.uk/thc/](http://www.infowar.co.uk/thc/) toneloc

AT&T Information Security Center [www.att.com/isc/](http://www.att.com/isc/) docs/war\_dialer\_detection.pdf

## References

---

1. Information Assurance Technical Framework (IATF) <http://www.iatf.net> .
2. National Institute of Standards and Technology <http://niap.nist.gov/>.

## Additional References

---

- a. Amoroso, Edward, Intrusion Detection. Intrusion.Net Books. 1999.
- b. Escamilla, Terry. Intrusion Detection, Network Security Beyond the Firewall. Wiley Computer publishing. 1998.
- c. Northcutt, Stephen. Network Intrusion Detection, An Analyst's Handbook. New Riders Publishing. 1999.
- d. Concurrent Technologies Corporation. Attack Sensing, Warning, and Response (ASW&R) Baseline Tool Assessment Task War Dialer Trade Study Report. Report No. 0017-UU-TS-000480. March 23, 2000.
- e. King, Nathan A. Sweeping Changes for Modem Security. Information Security Magazine. Volume 3, Number 6. June 2000.
- f. Ulsch, Macdonnell and Joseph Judge. Bitter-Suite Security. Information Security Magazine. Volume 2, Number 1. January 1999.
- g. Department of Defense (DoD) Chief Information Officer (CIO) Guidance and Policy Memorandum No. 6-8510, Guidance and Policy for Department of Defense Global Information Grid Information Assurance.
- h. National Security Telecommunications and Information Systems Security Policy (NSTISSP) No. 11. National Policy Governing the Acquisition of Information Assurance (IA) and IA-Enabled Information Technology (IT) Products. January 2000.

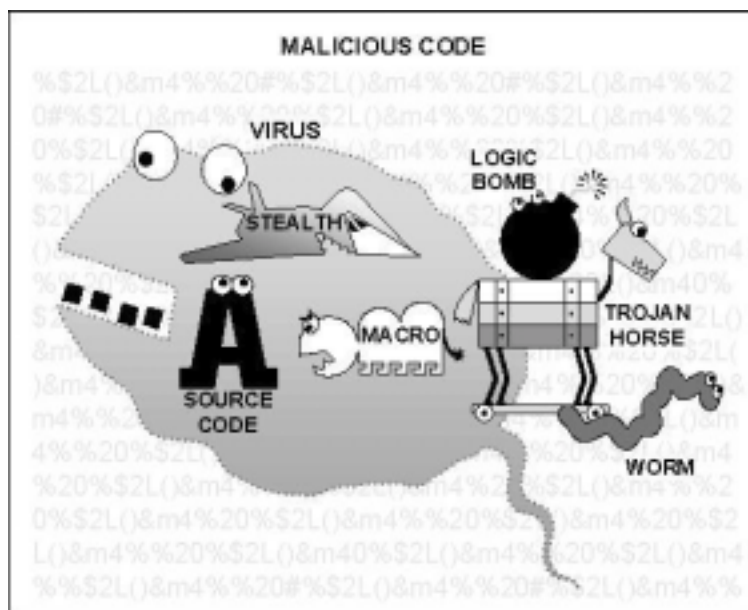


## 6.6 Malicious Code Protection

The objective in this section of the Framework is to elucidate the importance of defense from destructive malicious code. Information is provided regarding malicious code protection techniques and how malicious code infiltrates a system. Detection and recovery tactics are described as well as different types of malicious code scanners used to protect systems.

Malicious code protection allows authorized Local Area Network (LAN) users, administrators, and individual workstation/ personal computer users to safely conduct daily functions in a secure manner. Commonly, many people misuse the word virus assuming it means anything that infects your computer and causes damage. This term is really referred to as malicious code. A virus is simply a computer program created to infect other systems/ programs with copies of itself. Worms are similar to viruses; however, they do not replicate and the intent is usually destruction. Logic Bombs contain all types of malicious code and activate when certain conditions are met. Viruses, worms, and logic bombs can also be concealed within source code disguised as innocent programs like graphic displays and games. These programs describe a Trojan Horse. The relationship among these different types of malicious code is illustrated in Figure 6.6-1 below.

The quantity of new malicious code introduced into the computing environment has increased exponentially. This situation has occurred for several reasons. Computer users have become increasingly proficient and sophisticated and software applications have become increasingly complex. Some brands of software are now widely used, thus their bugs and security loopholes are often known to intelligent users capable of writing destructive code. With the widespread use of personal computers that lack effective malicious code protection mechanisms, it is relatively easy for knowledgeable users to author malicious software and dupe unsuspecting users into copying or downloading it. In addition, since virus information and source code is readily available through the Internet and other sources, creating viruses has become a relatively simple task.

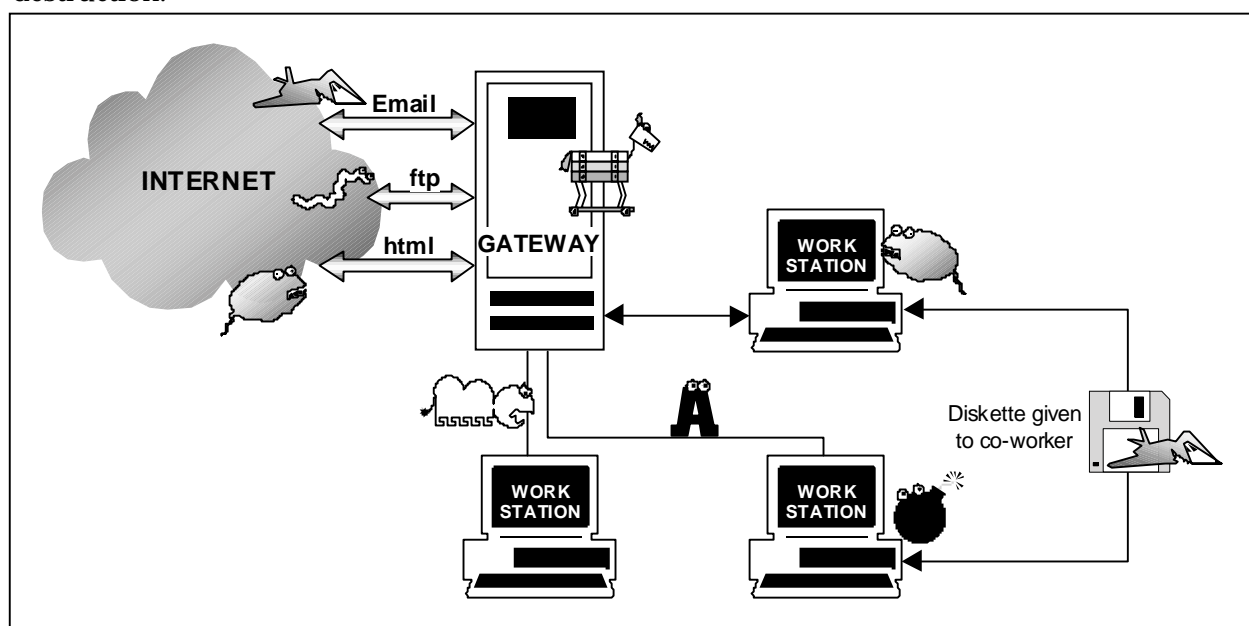


**Figure 6.6-1 Malicious Code Relationships**

### 6.6.1 Target Environment

Two generic cases exist: protection for networks at the gateway and protection for workstations with individual access to networks or information services. Malicious code can infiltrate and destroy data through network connections if allowed beyond the gateway or through

individual user workstations. Today, the majority of individual users keeps all data files on networks or shared file systems instead of on diskettes. Therefore, continual protection of network connections at the gateway is essential. Malicious code usually enters existing networks through the gateway by entering through security loopholes or e-mail attachments. Their intent is to cripple the network and individual workstations. Malicious code can also attack the network through protocols, typically, File Transfer Protocol (FTP), Hypertext Transfer Protocol (HTTP) and Simple Mail Transfer Protocol (SMTP) (e-mail). The individual user workstation is then subsequently infected. In Figure 6.6-2 below, a simplified network is illustrated with several workstations connected to a single gateway subsequently connected to the Internet. Although a single user can bring an infected disk to work, infecting his workstation and eventually the entire network, the majority of infections by malicious code result from file sharing across different protocols. Malicious codes attacking individual user workstations are primarily macro viruses and other potentially less destructive viruses. These viruses typically enter systems through e-mail attachments; however, their primary intent is not destruction.



**Figure 6.6-2. Sources of Malicious Code Infections**

## 6.6.2 Malicious Code Protection Requirements

### 6.6.2.1 Functional Requirements

#### Malicious Code Detection System Requirements

The following have been identified as representative malicious code detection system requirements from a customer's perspective of needs.

The malicious code detection system shall

- Allow access to all services available on the Wide Area Networks (WAN) using any of the existing and emerging networking technologies and applications.
- Be able to locate the source and type of an infection, be able to react to such intrusions, and be able to fully reconstitute the system following damage caused by intrusions.
- Have minimal operational effect on the user.
- Have minimal operational effect on performance of the associated components.
- Have appropriate documentation for its use and upgradability and shall contain all currently available references and resources.
- Allow automatic malicious code prevention programs to run in the background.
- Allow a disaster recovery plan to recover data if necessary.
- Provide adequate scanning tools to be able to contain an identified virus by isolating affected systems and media.
- Have appropriate means to trace all incoming and outgoing data, including e-mail, FTP transactions, and Web information.
- Be able to, in the event the Internet is unavailable for any reason, still have access to virus updates from the manufacturer or vendor of the anti-virus product.
- Monitor usage as required by the administrator.
- Scan for malicious software at the enclave boundary and at individual workstations.
- Log and analyze source-routed and other packets; react to or restrict malicious code attacks.
- Allow a rapid disconnect from the network in the event of a detected malicious code attack.

## Configuration/ Management Requirements

The following have been identified as representative configuration and/or management requirements for malicious code detection systems.

The malicious code detection system shall:

- Be updated with regard to relevant security issues (malicious code detection, system vulnerability) so maximum protection is provided.
- Be capable of preventing worm programs from infecting networks by allowing the administrator to disable the network mail facility from transferring executable files.
- Be configured by the administrator to filter all incoming data including e-mail, FTP transactions, and Web information for all types of malicious code.
- Allow the administrator to automatically create policy for network usage that details what sort of computing activity will and will not be tolerated.
- Allow regular backups of all system data by the administrator.
- Provide adequate controls such as strong user authentication and access control mechanisms on network connections for the administrator.

- Be capable of setting additional passwords or authentication for select files and accounts accessed from network ports.
- Be capable of placing restrictions on types of commands used on networks and in select files.
- Not permit access to system manager accounts from network ports, if possible.
- Monitor usage of the network during odd hours, if possible and create a log of all activity for the system administrator.
- Not provide more than one administrator account, i.e., shall not give other users administrator privileges.

## 6.6.3 Potential Attack Mechanisms

Malicious code can attack authorized LAN users, administrators, and individual workstation/personal computer users in numerous ways such as modifying data in transit, replaying (Insertion of Data), exploiting data execution, inserting and exploiting malicious code, exploiting protocols or infrastructure bugs and modifying malicious software during production and/ or distribution. (See Sections 4.2.1.4.2 Network-Based Vulnerabilities and Active Attacks, and 4.2.1.4.4 Hardware/Software Distribution.)

### 6.6.3.1 Viruses and Worms

The operating system (OS) is software that controls all inputs and outputs to the system and manages the execution of programs. A virus or worm can infect the OS in two ways: by completely replacing one or more OS programs or by attaching itself to existing OS programs and altering functionality. Once a virus or worm has altered or changed OS functionality, it can control many OS processes that are running. To avoid detection, the virus or worm usually creates several hidden files within the OS source code or in “unusable” sectors. Since infections in the OS are difficult to detect, they have deadly consequences on systems relying on the OS for basic functions.

#### Macro Viruses

Application programs on a system provide users with significant functionality. A macro virus can easily infect many types of applications such as Microsoft Word and Excel. To infect the system, these macro viruses attach themselves to the application initialization sequence. When an application is executed, the virus' instructions execute before control is given to the application. These macro viruses move from system to system through e-mail file sharing, demonstrations, data sharing and disk sharing. Viruses that infect application programs are the most common and can lie dormant for a long time before activating. Meanwhile, the virus replicates itself, infecting more and more of the system.

### 6.6.3.2 Logic Bombs

After a logic bomb has been activated, it can maliciously attack a system in the following ways: halt machine, make garbled noise, alter video display, destroy data on disk, exploit hardware

defects, cause disk failure, slow down or disable OS. It can also monitor failures by writing illegal values to control ports of video cards, cause keyboard failure, corrupt disks and release more logic bombs and/or viruses (indirect attacks). These attacks make logic bombs an extremely destructive type of malicious code.

### 6.6.3.3 Trojan Horses

Trojan Horses are another threat to computer systems. Trojans can be in the guise of anything a user might find desirable, such as a free game, mp3 song, or other application. They are typically downloaded via HTTP or FTP. Once these programs are executed, a virus, worm, or other type of malicious code hidden in the Trojan Horse program is released to attack the individual user workstation and subsequently a network.

### 6.6.3.4 Network Attacks

With the number of networks increasing exponentially, potential threats to these networks are numerous and devastating. The most common attack is to deny service by generating large volumes of Transmission Control Protocol/Internet Protocol (TCP/IP) traffic. The target site is rendered “unavailable” to the rest of the Internet community. The next level of denial of service (DoS) attacks is the distributed denial of service (DoS) attack where several machines on the target site are exploited. Distributed DoS attacks are the most effective and insidious because they generate more traffic from other sources, making it much harder to identify the attacker's source, and subsequently more difficult to resolve. An example of a distributed DoS attack was the attack by “coolio” in February 2000 which caused the crash of numerous Web sites in the United States (U.S.) including Ebay, CNN, Yahoo! , E\* Trade. This attack involved sending Internet Control Message Protocol (ICMP) echo request datagrams (ping packets) to the broadcast address of networks using a faked or “spoofed” IP address of the host to be attacked. The IP host responds to these ICMP echo requests on either the nominal address or the broadcast address of its interfaces. When the broadcast address of a network was pinged, all active hosts on that network responded and for any one request, there were many replies. This amplification makes distributed DoS attacks very powerful and causes large networks to crash.

### 6.6.3.5 Trapdoors

Access through trapdoors provides easy access for system administrators and authorized personnel to gain entry to a system or a system's resources. They can usually do this without a password. When these trapdoors are exploited, however, threats to a computer system are created. Authorized or unauthorized users with knowledge of trapdoors, can plant various types of malicious code into sensitive areas of a system. Therefore, the first layer of defense, prevention of malicious code, is bypassed, and the system must rely on detection and removal mechanisms to rid the system of the newly introduced malicious code.

### 6.6.3.6 Insider Attacks

Traditionally, insiders are a primary threat to computer systems. Insiders have legitimate access to the system and usually have specific goals and objectives. They can affect availability of system resources by overloading processing or storage capacity, or by causing the system to crash. Insiders can plant Trojan Horses in sensitive data files, which attack the integrity of the entire system. Insiders can also exploit bugs in the OS by planting Logic Bombs or by causing systems to crash. All of these attacks by insiders are difficult to prevent, as legitimate access is essential to all users for crucial daily functions.

### 6.6.3.7 Connection/Password Sniffing

Other threats to the integrity of a system include connection and password “sniffing.” A “sniffer” is malicious software or hardware that monitors all network traffic, unlike a standard network station that only monitors network traffic sent explicitly to it. Software sniffers can be a real threat to a network because they are “invisible” and easily fit on all workstations and servers. The specific threat presented by sniffers is their ability to catch all network traffic, including passwords or other sensitive information sent in plain text. An added threat to network security is that detecting sniffers on other machines is extremely difficult.

## 6.6.4 Potential Countermeasures

This section is subdivided into six types of countermeasures that can be applied to prevent and/or remove malicious code: Malicious code scanning products, electronic security (access constraint countermeasures), trapdoor access constraints, network security, connection and password sniffing countermeasures, and physical security.

### 6.6.4.1 Malicious Code Scanning Products

Malicious code scanning products are used to prevent and/or remove most types of malicious code including viruses, worms, logic bombs and Trojan Horses from a system. The use of malicious code scanning products with current virus definitions is crucial in preventing and/or detecting attacks by all types of malicious code.

### 6.6.4.2 Electronic Security

Electronic security typically refers to access constraint mechanisms used to prevent malicious code from infiltrating a system intentionally or unintentionally by authorized users. Unintentional system infiltration is the primary reason to implement access constraint mechanisms. If a set number of attempts to input a password correctly is exceeded, the system administrator must be contacted immediately. The system or system administrator should ensure that users change their passwords frequently and should not allow the use of “dictionary” words. This prevents against easy decryption of passwords. Checksums can also be used; however, they only pertain to some strains of viruses. All of these electronic security

measures protect against unintentional or angry employees deploying malicious code into a system or network.

The following are additional access constraint countermeasure requirements:

- Provide data separation – For data that is allowed access to the protected network workstation, steps should be taken to constrain the portion of the system that can be affected in case of a malicious code attack.
- Employ application-level access control – Access restrictions may also be implemented within a workstation or at various points within a LAN to provide additional layers and granularity of protection against authorized and unauthorized malicious code attacks.

### 6.6.4.3 Trapdoor Access/Distribution

To protect against unauthorized use of trapdoors regarding malicious code, reliable companies should be used when considering software and hardware purchases. When inputting data, only use reliable inputting individuals and use monitoring devices to monitor them. Reliable system administrators should remove passwords immediately after an employee leaves a company. All of these prevention techniques are crucial to prevent malicious code from infiltrating systems through trapdoors.

### 6.6.4.4 Network Security

A boundary protection mechanism at the gateway must be used within a network. The requirements for a boundary protection mechanism are mentioned in the following sections of the Information Assurance Technology Framework (IATF): firewalls, guards, intrusion detection, WAN. The requirements in these sections describe a boundary protection mechanism for network security.

There are also several ways to prevent a network against distributed DoS attacks by malicious code. Secure hosts on the network by replacing “rlogin” and “rexec” commands with “ssh”, or other encrypted commands. Also, disallow IP spoofing to keep hosts from pretending to be others. Do not allow ICMP to broadcast and multicast addresses from outside the network. These few preventative methods will help prevent distributed DoS attacks.

### 6.6.4.5 Connection and Password Sniffing Countermeasures

Although “sniffing” of Internet traffic is difficult to stop, there are several ways to defend a system and make “sniffing” difficult. First, use an encryption mechanism (e.g., Secure Socket Layer (SSL)) to allow encryption of message transmissions across Internet protocols whenever possible. Also, encrypt e-mail through the use of Pretty Good Privacy (PGP) and Secure Multi-Purpose Internet Mail Extensions (S/MIME). Although e-mail is sent encrypted, when e-mail is read it must be unencrypted. If mail programs allow attachments to automatically run, malicious code can still infect a system. The malicious code will be encrypted with the rest of the message, and activate when you read the decrypted message. Also, implement “ssh”, or other encrypted commands instead of insecure remote login. To stop password “sniffers” use

secure remote access and smart cards to keep passwords private. To protect a LAN from “sniffing” replace a hub with a switch, which is extremely effective in practice. Although “sniffers” can still access the LAN, it becomes more difficult to do so.

### 6.6.4.6 Physical Security

To be physically secure against potential infections by malicious code, the system must be protected from physical attack. It is necessary to use a monitoring system to authenticate users to restrict physical access. Once access is granted, users' actions must be monitored.

### 6.6.4.7 Detection Mechanism

The detection mechanism enables users to detect the presence of malicious code, respond to their presence, and recover data or system files, if possible.

#### Detect

The objectives for detection are to discover attacks at or inside the protected boundary as well as to facilitate tracking and prosecuting of adversaries. Malicious code detection involves the continual probing of internal networks for the existence of services or applications infected by malicious code. This may be done routinely to assist in the selection of additional appropriate countermeasures, to determine the effectiveness of implemented countermeasures, or to detect all types of malicious code. The following are typical security capability requirements associated with malicious code detection and system probing.

- Provide centralized operation.
- Provide automated reports.
- Recommend corrective action.
- Archive significant security events.
- Display and record status in real time.

#### Respond

To respond to the presence of detected malicious code within a system or network, malicious code scanning must be performed. The following are typical security capability (countermeasure) requirements.

- Detect occurrence of infection and locate malicious software, e.g., virus, memory.
- Perform scanning automatically, e.g., run continual malicious code scans throughout the day on systems.
- Implement scanning at the network gateway and at network components such as the desktop.
- Identify specific malicious code, e.g., macro virus.
- Remove malicious code from all infected systems so it cannot infect further, e.g., boot from uninfected write-protected boot diskette, then remove the malicious code from the system.



- Correct all effects of malicious code and restore system to original state, e.g., check all diskettes with files that may have been in disk drives during virus residency; reload files as appropriate.
- Reload program backups in cases where malicious code cannot be completely identified or where removal is not possible.
- Perform manually initiated scanning regularly, e.g., scan for malicious code after any Internet downloads.

## Recover

To recover data from the infection of malicious code, first concentrate on the specific area infected. If malicious code infects many programs, concentrate on all programs for infection and recovery. The recovery process will take longer if malicious code has been in the system. The number of computers that have been infected is also important as it affects time and resources for recovery. There are four stages in the infection process, and each stage requires a different amount of time and resources for recovery.

- 1) Local Memory Infection is the first stage of the infection process of a malicious code. If malicious code is caught in the first few hours before an appropriate host is found and replication begins, the following straightforward approach can be applied:
  - a) Power down,
  - b) Cold reboot with a clean, write-protected diskette,
  - c) Run a utility program to check hard disk and remove the few infected files, and
  - d) Locate and destroy the source containing the malicious code.
- 2) Local Disk Storage Infection is the second stage of the infection process. If an infection goes undetected, malicious code will infect an increasing number of programs and data files over time. In this case, the removal process becomes more complicated and several things could happen. If data and program files have been destroyed, it is possible a reformat of the entire media would be required for recovery. File backups can also be dangerous due to the risk of re-infection during the restoring process. Total data loss may occur.
- 3) Shared File System Infection is the third stage of the infection process of malicious code. The risk of malicious code infecting the network attached to a computer is very high. If the infection is widespread, it is possible that a reformat of the entire media would be required for recovery. Many things could happen during the recovery process. Again, file backups can be dangerous due to the risk of re-infection during the restoration process. One complication is numerous computers attached to the infected network will also be infected. The malicious code must be removed simultaneously from all workstations as well as the network. Another complication is that other users may have saved the malicious code unknowingly onto a floppy disk that may infect the entire network later.
- 4) System-Wide Removable Media Infection is the final stage of the infection process. An infected computer will infect many of the physical disks it contacts. This is an extremely difficult situation to deal with for numerous reasons. Malicious code infects all types of

removable media such as floppy diskettes, removable hard disks, reel and cartridge tapes, etc. Once an infected disk has successfully infected a network computer, the number of infected disks drastically increases. A complication with all the infected disks is the possibility of re-infection after malicious code has been discovered and removed. Although scanning devices would have been updated since the original infection and would catch many possible re-infections, new malicious code like the polymorphic virus that changes itself after each infection could still compromise the network. Malicious code could also reach client sites and computers.

## 6.6.4.8 Administrative Countermeasures

Administrative concerns regarding infection due to malicious code include training, policy, and coping with fears about malicious code and computers. “Viruses affect the emotional relationships that many people develop with their computer. Viruses could change the very nature of computing, from an essentially logical, predictable function to one fraught with uncertainty and danger.” It is crucial for administrators to minimize stress fears due to computer viruses while not blaming employees.

Administrators can combat fears about malicious code and computers in many ways. The staff should be educated and motivated with regard to malicious code protection, detection, and recovery. A review of computer security with a risk analysis of exposure to infection and likely consequences should be conducted. A corporate policy with information about malicious code should be distributed to all staff. In addition, special briefing sessions should be held for all staff involved with computing functions. Administrators need to institute prevention programs that incorporate safe computing practices that should be posted at all terminals. Regular training sessions on safe computing should be scheduled. Administrators should also have a disaster recovery plan that is practiced on worst-case scenarios. Twenty-four hour emergency phone numbers should be displayed. Most employees should also be cautioned to avoid overreaction and deploy back-up facilities to minimize consequential damage.

## 6.6.5 Technology Assessment

Before describing malicious code detection products, it is important to understand different types of malicious code.

### 6.6.5.1 Types of Malicious Code

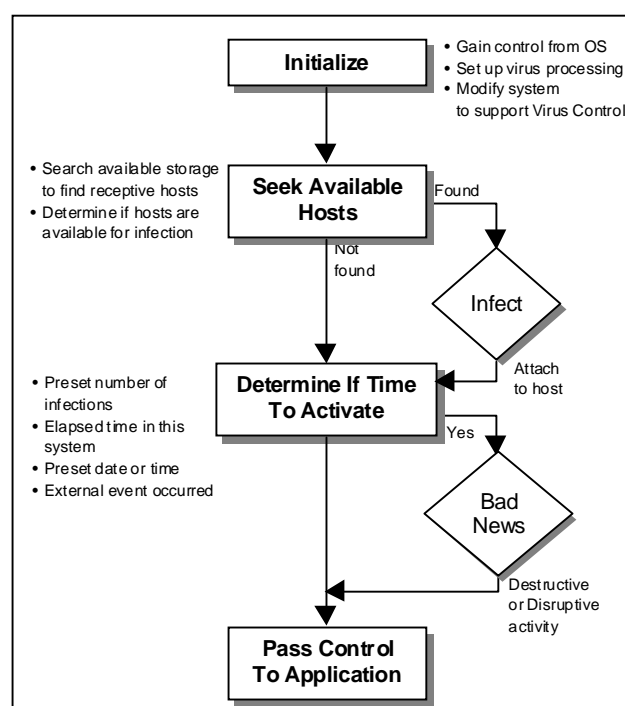
#### Viruses

Regarding malicious code, there are several classes of viruses that range from innocuous to catastrophic. An understanding of each class is crucial to understanding the evolutionary process of an infiltrating virus. Innocuous viruses reside in unobtrusive areas of the system and cause no noticeable disruption. These viruses infect diskettes and other media that come into contact with the system but intend no damage. Humorous viruses cause aggravating events to occur, humorous messages to appear, or graphic images to be displayed. Although irritating, these viruses intend no damage and are commonly used for jokes. Potentially the

most disruptive and difficult to detect are the data-altering viruses that alter system data. The viruses modify data file numeric information in spreadsheets, database systems, and other applications, such as changing all occurrences of the number three to the number eight. Catastrophic viruses erase critical system files and immediately cause widespread destruction. The viruses scramble key information tables and/or remove all information on all disks, including shared and network drives.

There are two main phases in the lifecycle of a virus.

- 1) The first phase, replication, could last a few weeks to several years. In this phase, viruses typically remain hidden and do not interfere with normal system functions. Viruses also actively seek out new hosts to infect such as attaching itself to other software programs or infiltrating the OS. A virus that is attached to an executable program executes its instructions before passing control to the program. (See Figure 6.6-3) These viruses are hard to detect because they only infect a small amount of programs on a disk and the user is not suspect.
- 2) During the second phase, activation, the beginning of gradual or sudden destruction of the system occurs. Typically, the decision to activate is based on a mathematical formula with criteria such as date, time, number of infected files, and others. The possible damage at this stage could include destroyed data, software or hardware conflicts, space consumption and abnormal behavior.



**Figure 6.6-3 Virus Execution**

Extraordinary damage caused by these viruses can result in loss of man-hours, disruption of normal activities and wasted monetary resources. Therefore, the unrealized potential for harm is the main reason why malicious code scanning and prevention is extremely important.

## Macro Viruses

The 1995 advent of macro programming for applications, like MS Word and Excel, automated repetitive keystroke functions, but also created an effective new way for viruses to spread. Word and Excel data files had previously been data-only files, like text-only e-mail messages – unable to harbor viruses because they did not include executable code.

Virus writers soon discovered these applications' macros could also be used to create viruses. At the same time, sharing of documents and spreadsheet files via e-mail became increasingly

commonplace between users both within and between companies – creating the most effective virus carrier ever. Among the factors contributing to the dominance of macro viruses is the Visual BASIC for Applications (VBA) programming language that makes it as easy for users to create legitimate timesaving macro commands as it does for virus writers to create time-robbing macro viruses.

Once the macro-infected file is accessed, it replaces one of the Word or Excel standard macros with an infected version that can then infect all subsequent documents it comes into contact with. Macro viruses usually disable the macro menu selection, making users unable to see what macros are executing.

Today, macro viruses like ILOVEYOU are the most prevalent computer viruses in the wild – accounting for the vast majority of virus encounters in corporations. Today's widespread sharing of macro-enabled files, primarily through e-mail attachments, is rapidly increasing along with the associated macro virus threat.

The Table 6.5-1 Comparison of Macro Viruses describes the current impact of several macro viruses compared to an older virus and the associated costs to corporations.

**Table 6.6-1 Comparison of Macro Viruses**

<b>Virus</b>	<b>Year</b>	<b>Type</b>	<b>Time to become prevalent</b>	<b>Estimated Damages</b>
Jerusalem, Cascade, Form	1990	Executable file, boot sector	Three Years	\$50 million for all viruses over five years
Concept	1995	Word macro	Four months	\$60 million
Melissa	1999	E-mail enabled Word macro	Four days	\$93 million to \$385 million
I Love You	2000	E-mail enabled Visual Basic script/ word macro	Five hours	\$700 million

## Polymorphic Viruses

Polymorphic viruses alter their appearance after each infection. Such viruses are usually difficult to detect because they hide themselves from anti-virus software. Polymorphic viruses alter their encryption algorithm with each new infection. Some polymorphic viruses can assume over two billion different guises. This means anti-virus software products must perform heuristic analysis, as opposed to spectral analysis that can find simpler viruses.

There are three main components of a polymorphic virus: a scrambled virus body, decryption routine, and a mutation engine. In a polymorphic virus, the mutation engine and virus body are both encrypted. When a user runs a program infected with a polymorphic virus, the decryption routine first gains control of the computer, then decrypts both the virus body and the mutation engine. Next, the decryption routine transfers control of the computer to the virus, which locates a new program to infect. At this point, the virus makes a copy of itself and the mutation engine in random access memory (RAM). The virus then invokes the mutation engine, which randomly generates a new decryption routine capable of decrypting the virus,

yet bears little or no resemblance to any prior decryption routine. Next, the virus encrypts the new copy of the virus body and mutation engine. Finally, the virus appends the new decryption routine, along with the newly encrypted virus and mutation engine, onto a new program. As a result, not only is the virus body encrypted, but also the virus decryption routine varies from infection to infection. This confuses a virus scanner searching for the telltale sequence of bytes that identifies a specific decryption routine. Therefore, with no fixed signature to scan for, and no fixed decryption routine, no two infections look alike.

A good way to contain a polymorphic virus is to set up false data directories or repositories to fool the attacker into thinking they have reached exploitable data. This can significantly reduce the risk of being attacked. The polymorphic virus executes in these false data directories, and is fooled into believing it has infected the entire system. In reality, the directories are either deleted or non-existent and the virus is thus unable to infect the system.

## Stealth Viruses

Stealth viruses attempt to hide their presence from both the OS and the anti-virus software. Some simple techniques include hiding the change in date and time as well as hiding the increase in file size. Stealth viruses sometimes encrypt themselves to make detection even harder. Stealth viruses also enter systems through simple download procedures. Unsuspecting users can do little against this type of infection except download files only from trusted sources.

## Worms

Worms are constructed to infiltrate legitimate data processing programs and alter or destroy the data. Although worms do not replicate themselves as viruses do, the resulting damage caused by a worm attack can be just as serious as a virus, especially if not discovered in time. However, once the worm invasion is discovered, recovery is much easier because there is only a single copy of the worm program to destroy since the replicating ability of the virus is absent.

A prevalent worm, "Ska" is a Windows e-mail and newsgroup worm. An e-mail attachment disguised as "Happy99.exe" will display fireworks when executed the first time. After execution, every e-mail and newsgroup posting sent from the machine will cause a second message to be sent. Since people receive "Happy99.exe" from someone they know, people tend to trust this attachment, and run it. Then the worm causes damage by altering functionality of the WSOCK32 dynamic library link (DLL) file. Now the worm can actively attack other users on the network by placing itself on the same newsgroups or same e-mail addresses where the user was posting or mailing to.

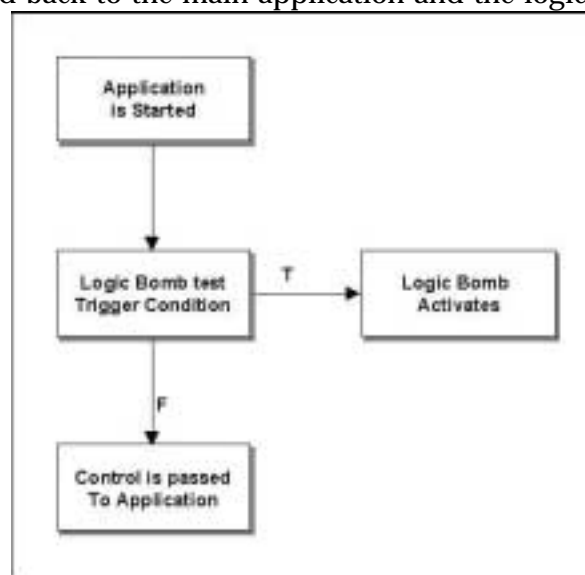
## Trojan Horses

Trojan Horse is an apparently harmless program or executable file, often in the form of an e-mail message that contains malicious code. Once a Trojan gets into a computer or network, it can unleash a virus or other malicious code, take control of the computer infrastructure and compromise data or inflict other damage. The Melissa virus that struck in 1999 is a good example of a harmful Trojan. Attached to a harmless-looking e-mail message, the virus accessed Microsoft Outlook, replicated itself, and sent itself to many other users listed in the recipient's e-mail address book. The resulting e-mail-sending flurry caused many Microsoft Exchange servers to shut down while users' mailboxes flooded with bogus messages.

Trojans can also be carried via Internet traffic such as FTP downloads or downloadable applets from Web sites. These can not only compromise enterprise computers and networks by rapidly infecting entire networks, but they can also invite unauthorized access to applications that results in downtime and the cost to business could potentially reach into the millions of dollars.

## Logic Bombs

Logic bombs are programs added to an already existing application. Most are added to the beginning of the application it is infecting so it is run every time that application is run. When the infected program is run the logic bomb is run first and usually checks the condition to see if it is time to run the bomb. If not, control is passed back to the main application and the logic bomb silently waits (See Figure 6.6-4). When the right time does come, the rest of the logic bombs' code is executed. At that time the hard disk may be formatted, a disk erased, memory corrupted or anything else. There are numerous ways to trigger logic bombs: counter triggers, time triggers, replication triggers (activate after a set number of virus reproductions), disk space triggers, video mode triggers (activate when video is in a set mode or changes from set modes). Also, Basic Input Output System (BIOS) read only memory (ROM) triggers (activate when a set version of BIOS is active), keyboard triggers, anti-virus triggers (activate when a virus detects variables declared by virus-protection software such as "SCAN\_STRING") and processor triggers (activate if a program is run on a particular processor).



**Figure 6.6-4 Logic Bomb Execution**

Logic bombs cannot replicate themselves and therefore cannot infect other programs. However, if the program that is infected is given to someone else and the right conditions are met on that computer it will go off.

### 6.6.5.2 Viruses and E-mail

Today's office worker receives an average of more than forty e-mail messages each day. Many of these messages have Microsoft Word or Excel data files attached, that may carry macro viruses. Since plain text data cannot carry the executable program code viruses need to copy and spread themselves, the text messages of electronic mail are, by themselves, unable to spread viruses. The virus danger from e-mail stems from attachments containing active executable program files with extensions such as: CLASS, OCX, EXE, COM and DLL – and from macro-enabled data files. These attachments do not even need to be opened as many mail clients automatically display all attachments. To prevent attachments from automatically being displayed, simply configure the mail client to prompt the user. Another safeguard is to identify file extensions prior to opening attachments so the infection of many computer systems may be prevented. These attachments could contain malicious code that could be masquerading as another file type.

### 6.6.5.3 Virus Creation

There are two types of viruses that can be created: simple viruses and complex viruses.

#### Simple Viruses

Simple viruses do not attempt to hide themselves and are easy to write. Users with little computer knowledge can use Internet programs to create these viruses. Since thousands of sites contain virus source code, users can easily download and use existing viruses to infect systems. Users with slightly more computer knowledge may even alter existing virus source code or combine several viruses to create a new undetectable virus capable of compromising systems.

#### Complex viruses

Complex viruses require more source code than simple viruses used to conceal themselves from systems. Knowledge of assembly language is required to manipulate interrupts so these viruses can remain hidden. While hiding, complex viruses replicate, and will destroy data later. A complex virus is divided into three parts: the replicator, the concealer, and the bomb. The replicator part controls spreading the virus to other files, the concealer keeps the virus from being detected and the bomb executes when the activation conditions of the virus are satisfied. After these parts are created and put together, the virus creator can infect systems with a virus that current anti-virus software cannot detect.

### 6.6.5.4 Virus Hoaxes

The Internet is constantly being flooded with information about malicious code. However, interspersed among real virus notices are computer virus hoaxes. Virus hoaxes are false reports about non-existent viruses, often claiming to do impossible things. While these hoaxes do not infect systems, they are still time consuming and costly to handle. Corporations usually spend much more time handling virus hoaxes than handling real virus incidents. The most prevalent virus hoax today is the “Good Times Hoax” that claims to put your computer's central processing unit (CPU) in an “ $n^{\text{th}}$ -complexity infinite binary loop that can severely damage the processor.” In this case, there is no such thing as an  $n^{\text{th}}$ -complexity infinite binary loop. It is estimated virus hoaxes cost more than genuine virus incidents. No anti-virus product will detect hoaxes because they are not viruses, and many panic when they receive a hoax virus warning and assume the worst – making the situation much worse.

### 6.6.5.5 System Backup

There are two main strategies to follow when performing a system backup.

#### Workstation Strategy

The best backup strategy for workstations is to backup often. If the workstation is running the Windows OS, there are some simple backup tools already provided. There are also several utilities and programs available from other companies to assist users in performing backups.

The following features can make backup chores more bearable: incremental backup, unattended scheduling, and easy, simple restoration. Incremental backup saves changes made since the most recent full or incremental backup. This is important because users who do not want to wait to backup a system can use incremental backup as a substitute for a lengthy full backup. Scheduling uses software automation to execute backup chores without the need for personal interaction. Although a backup media must be selected and in place, the user does not need to be present for the actual backup. Zip drives and small tape drives are also cost-effective solutions used to backup workstation data.

## Network Strategy

The best backup strategy for networks is an approach that combines several features to save time, effort, and still assure complete backups. Execute full backups often. Since backups take up network, server, and/or workstation resources, it is best to run full backups when nobody is working. In addition, open files are skipped during backup and do not get backed up at all until some future time when the file is closed and not being used. Having few to no users holding files open will ensure the greatest backup saturation possible. Full backups are most efficiently executed in the evenings. Store the full backup tape off site. On each of the remaining workdays of the week, using a separate tape for each day, run an incremental backup and store it offsite, too. The last full backup of the month should be permanently moved off site and held for archival purposes. Therefore, if a network is attacked by malicious code, these backup techniques will ensure data integrity and allow all systems to be recovered.

## 6.6.5.6 Types of Malicious Code Detection Products

Most computer malicious code scanners use pattern-matching algorithms that can scan for many different signatures at the same time. Malicious code detection technologies have to include scanning capabilities that detect known and unknown worms and Trojan horses. Most anti-virus products search hard disks for viruses, detect and remove any that are found, and include an auto-update feature that enables the program to download profiles of new viruses so that it will have the profiles necessary for scanning. The viral signatures these programs recognize are quite short: typically 16 to 30 bytes out of the several thousand that make up a complete virus. It is more efficient to recognize a small fragment than to verify the presence of an entire virus, and a single signature may be common to many different viruses.

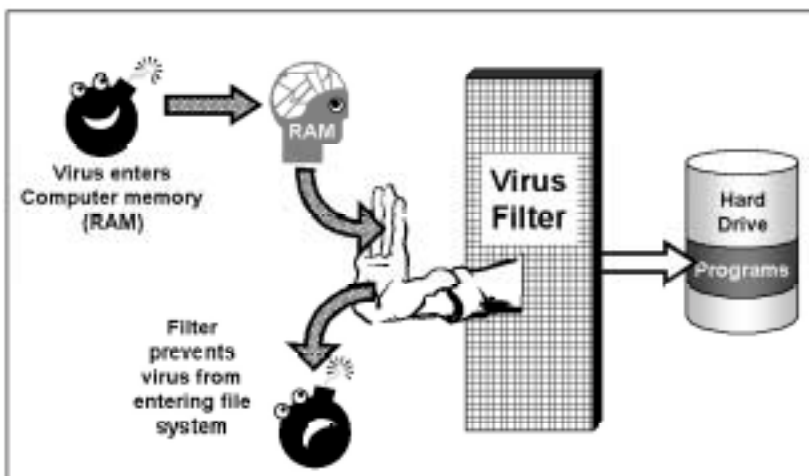
### 6.6.5.6.1 Pre-Infection Prevention Products

Pre-infection prevention products are used as the first level of defense against malicious code. Before they actually attack a system, prevention products should be applied. E-mail filtering products are available that do not allow executable programs or certain file types to be transferred. Also, options in browsers that limit the use of and/or disable Java and ActiveX plug-ins should be implemented. Simply changing browser options allows the user to see hidden files and file extension names. This could prevent opening an infected file masquerading as a normal text file. These essential pre-infection prevention products are the first level of defense against malicious code attacks.



### 6.6.5.6.2 Infection Prevention Products

Infection prevention products are used to stop the replication processes and prevent malicious code from initially infecting the system. These types of products, protecting against all types of malicious code, reside in memory all the time while monitoring system activity. When an illegal access of a program or the boot sector occurs, the system is halted and the user is prompted to remove the particular type of malicious code. These products act like filters that prevent malicious code from infecting file systems. (See Figure 6.6-5.)



### 6.6.5.6.3 Short-Term Infection Detection Products

Figure 6.6-5 Virus Filter

Short-term infection detection products detect an infection very soon after the infection has occurred. Generally, the specific infected area of the system is small and immediately identified. These products also detect all types of malicious code and work on the principle that all types of malicious code leave traces. Short-term infection detection products can be implemented through vaccination programs and the snapshot technique.

#### Vaccination Programs

Vaccination programs modify application programs to allow for a self-test mechanism within each program. If the sequence of that program is altered, a virus is assumed and a message is displayed. The drawbacks to this implementation include the fact that the boot segment is very hard to vaccinate, and the malicious code may gain control before the vaccination program can warn the user. The majority of short-term infection detection products use vaccination because it is easier to implement.

#### Snapshot Technique

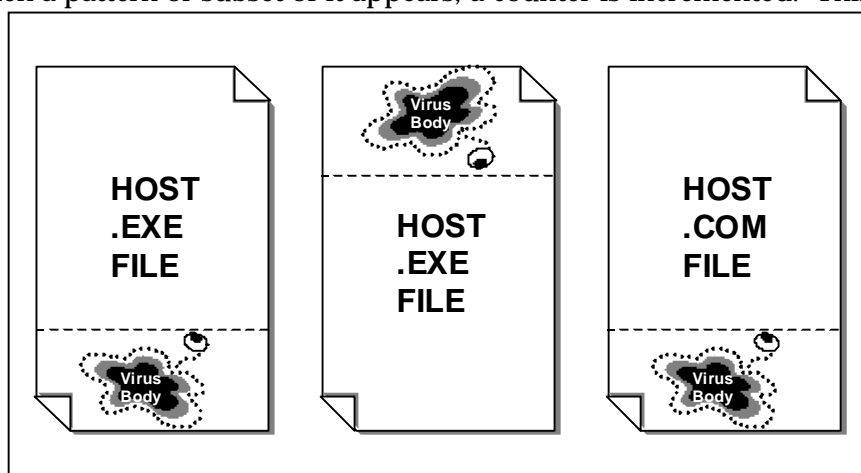
The snapshot technique has been shown to be the most effective. Upon installation, a log of all critical information is made. During routine system inspections (snapshots) the user is prompted for appropriate action if any traces of malicious code are found. Typically, these system inspections occur when the system changes: disk insertion, connection to different Web site, etc. This technique is difficult to implement in short-term infection detection products and is not widely used. However, when the snapshot technique is used with vaccination programs, an effective protection against malicious code is established.

### 6.6.5.6.4 Long-Term Infection Detection Products

Long-term infection detection products identify specific malicious code on a system that has already been infected for some time. They usually remove the malicious code and return the system to its prior functionality. These products seek a particular virus, and remove all instances of it. There are two different techniques used by long-term infection detection products: spectral analysis and heuristic analysis.

#### Spectral Analysis

Using spectral analysis, long-term infection detection products search for patterns from code trails that malicious code leaves. To discover this automatically generated code, all data is examined and recorded. When a pattern or subset of it appears, a counter is incremented. This counter is used to determine how often a pattern occurs. Using these patterns and the quantity of their occurrence, these products then judge the possible existence of malicious code and remove all instances of it. These products search for irregularities in code and recognize them as particular instances of malicious code.



**Figure 6.6-6 DOS File Infection**

#### Heuristic Analysis

Using heuristic analysis, long-term infection detection products analyze code to figure out the capability of malicious code. The underlying principle that governs heuristic analysis is that new malicious code must be identified before it can be detected and subsequently removed. This technique is much less scientific, as educated guesses are created. Because they are guesses, heuristic analysis does not guarantee optimal or even feasible results. However, it is impossible to scientifically analyze each part of all source code. Not only is this unproductive, it is terribly inefficient. Typically, good educated guesses are all that is needed to correctly identify malicious code in source code. These long-term infection detection products then remove all instances of the detected malicious code.

DOS file viruses typically append themselves on the end of DOS .EXE files. DOS file viruses prepend and append themselves onto DOS .COM files. (See Figure 6.6-6.) Other infection techniques are also possible but less common.

### 6.6.5.6.5 Interoperability

The different types of products mentioned above must be used together to create effective protection against all types of malicious code. Many layers of defense must be in place for a system to deal effectively with malicious code. If each type of product is implemented in a system, four different levels of defense are created. Before malicious code can attack a system, it must first get to the system through the prevention infection products. If it gets that far, the second layer of defense, prevention products will attempt to stop the malicious code from replicating. If that is not successful, then the detection products will try to locate and remove the infection before it reaches the majority of the system. Moreover, if the malicious code reaches the entire system, identification products can apply two different techniques to remove the infection. Each of these levels of defense is essential to the prevention and protection of a system.

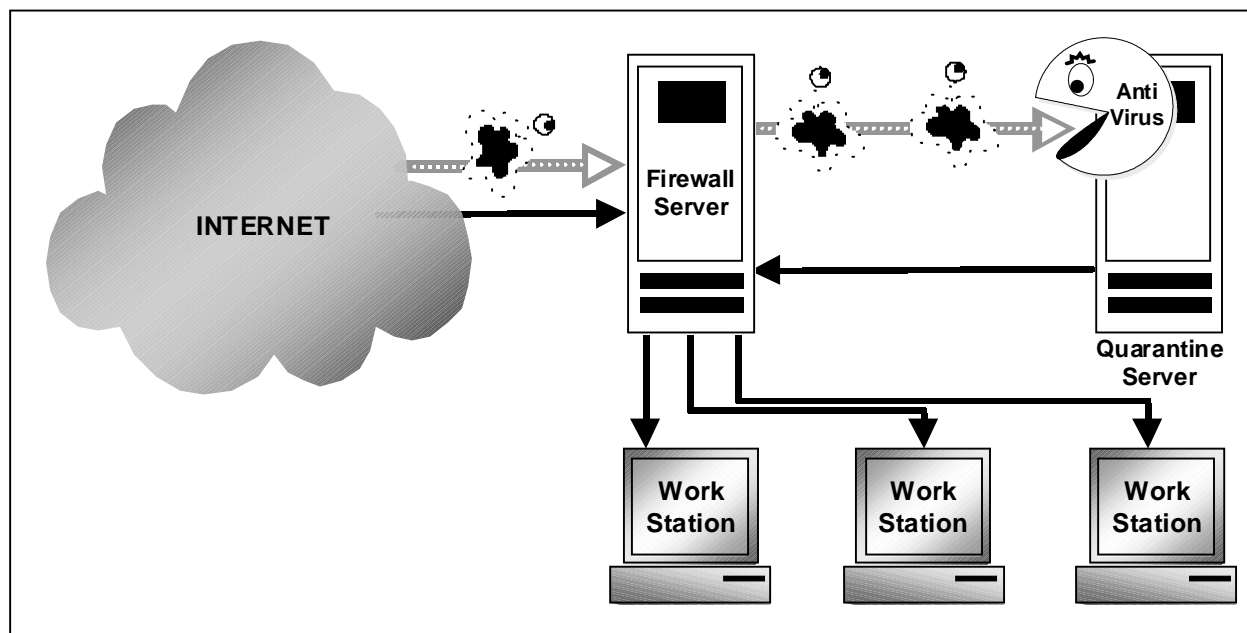
Today, commercial software packages combine all the above levels of defense and provide malicious code protection services. With new computer systems connecting to the Internet daily, security problems will also grow at an exponential rate. Unless a well-defined security policy is in place, IT managers will continue to lose the battle against computer viruses. Computer Emergency Response Team (CERT) statistics show the number of virus attacks rose from 3,734 in 1998 to 9,859 in 1999. In the first quarter of 2000 the Computer Emergency Response Team (CERT) has reported 4266 incidents. Despite the fact that anti-virus applications are essential for the detection of known viruses, no mail filter or malicious code scanner can defend against a new mail worm attack. The recent “Love Bug” virus was caught quickly and still did a wealth of damage. It seems to only be a matter of time before crackers figure out how to send e-mail worms that infect systems without opening attachments. While not sophisticated enough to stop new viruses from entering systems, anti-virus application makers are producing software that can prevent the damaging, data-altering affects of the malicious code.

### 6.6.5.7 Protection at the Workstation

There are numerous ways to protect a workstation from malicious code attacks. The implementation of pre-infection prevention, infection prevention, infection detection, and infection identification products provide four separate levels of defense and are essential in protecting a workstation. Although this is the best way to protect a workstation, other techniques can be applied. New malicious code protection products introduce a “sandbox” technology allowing users the option to run programs such as Java and ActiveX in quarantined sub-directories of systems. If malicious code is detected in a quarantined program, the system simply removes the associated files, protecting the rest of the system. Another protection mechanism is to allow continual virus definition updates that are transparent to the user. Implementing these updates at boot time, or periodically (1 hour, 2 hours, etc.) drastically reduces the chance a system will be infected with newly discovered malicious code. In the past six months alone, over 4,000 new viruses have been discovered. Without current viral definition updates, a system is left vulnerable to the devastating effects from malicious code.

## 6.6.5.8 Protection at the Network Gateway

When protecting a network, a number of issues must be considered. A common technique used in protecting networks is to use a firewall with Intelligent Scanning Architecture (ISA). (Figure 6.6-7) In this technique, if a user attempts to retrieve an infected program via FTP, HTTP, or SMTP, it is stopped at the quarantine server before it reaches the individual workstations. The firewall will only direct suspicious traffic to the anti-virus scanner on the quarantine server. This technique scales well since LAN administrators can add multiple firewall or gateway scanners to manage network traffic for improved performance. In addition, users cannot bypass this architecture, and LAN administrators do not need to configure clients at their workstations.



**Figure 6.6-7 Intelligent Scanning Architecture (ISA)**

Other useful scanning techniques for a network include continuous, automated malicious code scanning using numerous scripts. Simple commands can be executed and numerous computers in a network can be scanned for possible infections. Other scripts can be used to search for possible security holes through which future malicious code could attack the network. Only after fixing these security holes can a network withstand many attacks from malicious code.

## 6.6.6 Selection Criteria

When selecting anti-virus products, two important guidelines must be followed. The “best” product may not be good enough by itself. In addition, since data security products operate in different ways, one product may be more useful than another in different situations. When selecting a particular malicious code protection product, its installation must be considered. Is the program shipped on compact disk (CD), or on 1.44MB disks? Does the installation itself operate smoothly? There should be no questions without answers when properly installing a

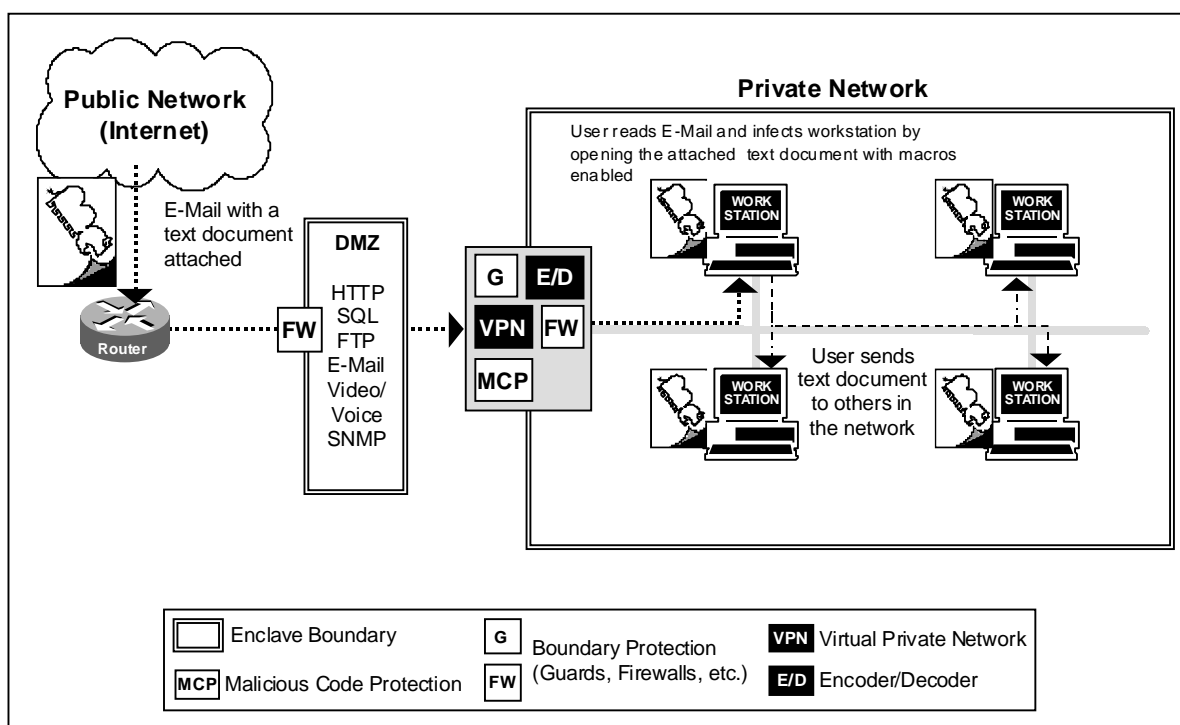
product. This product should be easy to use providing clear and uncluttered menu systems as well as meaningful screen messages.

Help systems are essential, as users need current information regarding all types of malicious code. The trend is to provide online help; however, manuals should also be provided with the product. The malicious code protection product should be compatible with all hardware and software and should not create conflicts. The company that produces the product should be stable and able to provide necessary local technical support for all questions and problems. The product should be fully documented. That is, all messages and error codes should be deciphered and full installation guides and how-to manuals should be provided. The computers to run this software must meet the hardware and software requirements specified by the manufacturer. The malicious code protection software should function properly and perform its duties without failing. Rating each of these categories will allow a company to choose the best malicious code protection product for their needs.

## 6.6.7 Cases

### 6.6.7.1 Case 1: Macro Virus Attack

Within a network environment, macro virus attacks are increasing exponentially. In Figure 6.6-8 below, a macro virus has infected an enclave via an e-mail attachment sent by an outsider. This e-mail attachment is a text document that enables macros. The e-mail recipient has e-mailed this document to his co-workers and saved it to diskette to view at home. A macro virus initiates when the document is opened and macros are enabled. As soon as the document is opened, the macro virus infects standard macros in the word processing program. After altering functionality of these standard macros, this virus replicates and infects many of the documents it comes into contact with.



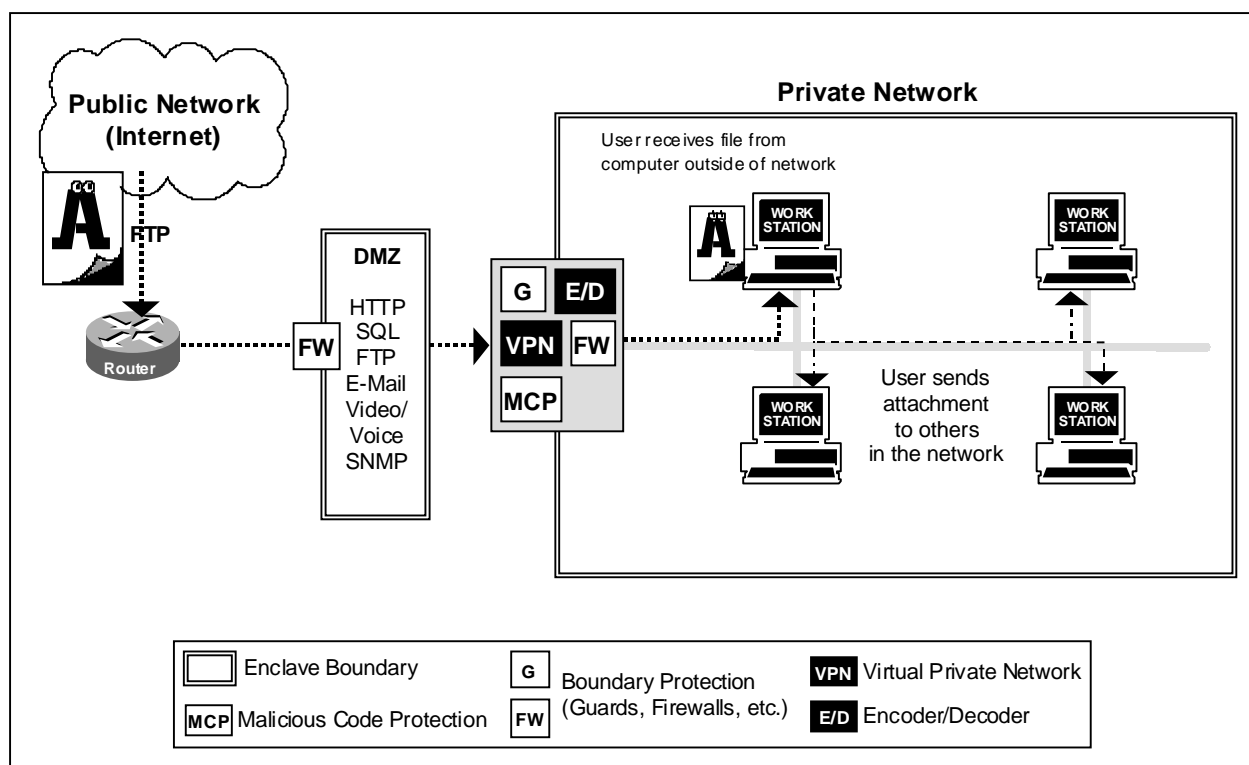
**Figure 6.6-8 Macro Virus Infection**

## 6.6.7.2 Case 2: Polymorphic Virus Attack

Polymorphic viruses represent the upper echelon of computer viruses. Today's polymorphic viruses are very difficult to detect using conventional anti-virus search engines because they possess the ability to mutate themselves and conceal their digital identity as they spread. The unique ability of this form of virus to change its signature to avoid detection makes it virtually undetectable, and therefore potentially disastrous in nature.

Polymorphic viruses infect enclaves in much the same way as macro viruses. In Figure 6.6-9 below, a polymorphic virus enters a system through FTP, as an unsuspecting user retrieves a single file from a computer outside the network. The user then sends this file via an e-mail attachment to other coworkers throughout the network.

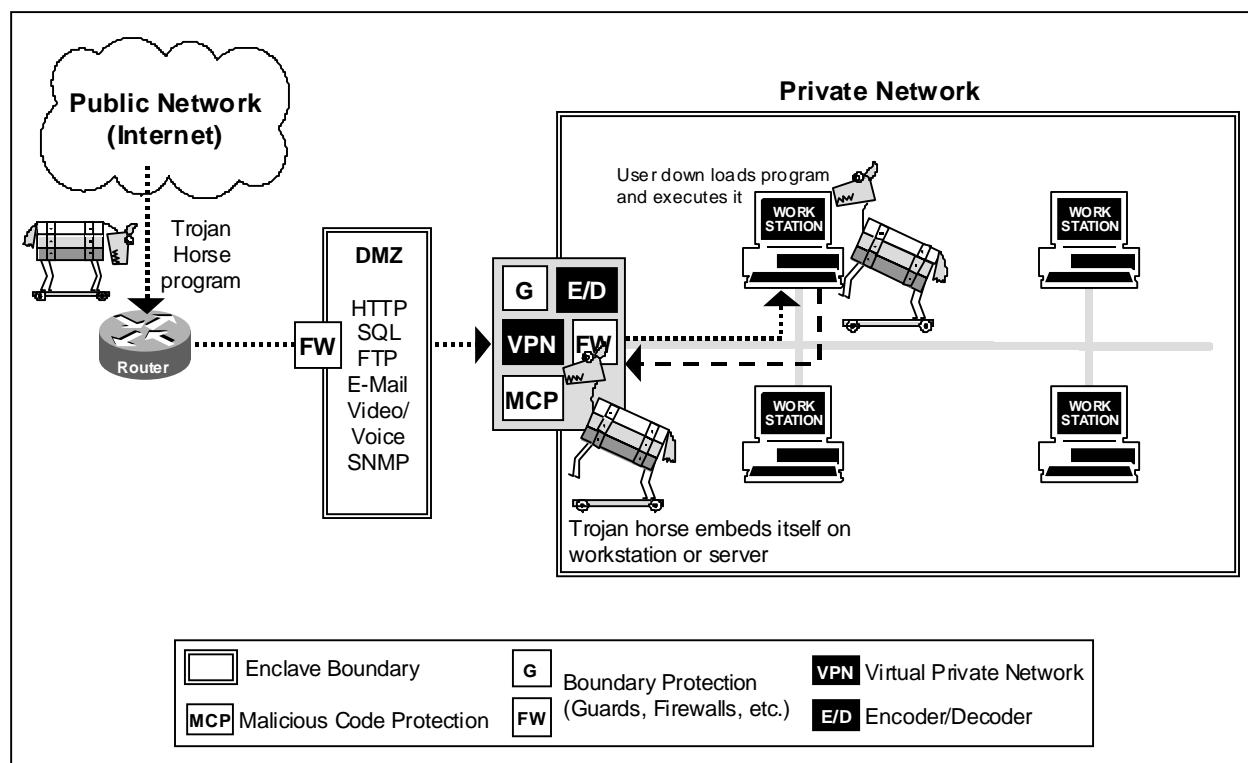
Once that file is accessed by any user, the polymorphic virus begins its programming and begins to replicate by e-mailing itself to the entire address book on its newfound host. It continuously changes its digital signature to escape the detection capabilities if any anti-viral application is resident.



**Figure 6.6-9 Polymorphic Virus Infection**

### 6.6.7.3 Case 3: Trojan Horse Attack

There exists a growing threat from another type of malicious software, the Trojan Horse. In Figure 6.6-10 below, a Trojan Horse has been imbedded into an existing network. A user downloaded a program that he thought was useful. However, after executing it, he realized it was not exactly what he needed. So, he deleted the file off of his computer. This unsuspecting user did not realize that the program downloaded was a Trojan Horse that imbedded itself into the network as a sniffer program after it was executed. Although this event occurred several weeks ago, there have been no problems in the network until now, when employees are noticing forged e-mails being sent to various clients.



**Figure 6.6-10 Trojan Horse Infection**

## 6.6.8 Framework Guidance

In this section guidance is provided on solutions that can be implemented so system infiltration by malicious code does not occur. Guidance will also be provided to detect and remove malicious code if it infects a system. Also, restoration guidance of the compromised system will be described.

### 6.6.8.1 Case 1: Macro Virus Attack

There are many ways to prevent, detect, respond and restore from macro virus attacks. The first level of defense is prevention so the macro virus does not reach the system. In a network environment, the first contact with the macro virus will be at the gateway. If the network is configured properly and using ISA (See Section 6.6.5.8 Protection at the Network Gateway.), the macro virus should be stopped at the quarantine server. It is crucial to have current virus definition updates in the malicious code detection software on the quarantine server. These updates should occur continually, and should be transparent to the user. Implementing these updates at boot time, or periodically (hourly) drastically reduces the chance a system will be infected by a newly discovered macro virus. So, these updates prevent new macro viruses from infecting the entire network. If the macro virus is not stopped at the gateway, individual workstations should detect the presence of the macro virus and remove them. At the next layer of defense, the individual user workstation will scan all incoming e-mail attachments for the presence of malicious code. If the malicious code detection software discovers the macro virus,



the file is simply deleted and the system and network are preserved. If virus updates are automatic, virus definitions for the quarantine server and the individual workstation should be the same, at the original time of system infiltration. In this case the detection software at the workstation will probably not detect the macro virus. If virus updates are not automatic, the individual user workstation will probably not detect the presence of the macro virus. This is because most users do not update their virus definitions as quickly as the system administrator of the quarantine server does. However, if this new macro virus has infected many workstations during a timeframe of several days, the possibility of vendors discovering this macro virus and updating their virus definitions increases. Once this macro virus is detected by an individual workstation, the system administrator should automatically be notified.

If the macro virus does infect the network by infecting workstations, the virus must be detected and removed. Typically, new macro viruses are detected when a user notices abnormal computer behavior and that abnormality is investigated. Another way to detect viruses is through automatic virus scanning with virus software definition updates. Once the presence of the macro virus is detected, it is essential to update all virus definition updates in all copies of malicious code protection software throughout the network. Then, several methods can be applied to remove all instances of the macro virus. If the infection has occurred recently (within a few hours), short-term infection detection products should be used. Using the snapshot technique, or vaccination programs, all instances of the macro virus are detected and then removed. If the infection is not recent, long-term infection detection products should be used. Using spectral and/or heuristic analysis, all instances of the macro virus are detected and then removed.

However, if the macro virus has fully infected network workstations, the macro virus removal will then allow for the data recovery process to begin. By practicing simple system backup procedures, (See Section 6.6.5.5 System Backup.) applications and data can be restored from tape backups with minimal data loss. After updating malicious code definitions for all malicious code protection software, the re-constituted network is then ready to proceed with daily functions. Any damage caused by the macro virus is removed and the system is restored to its prior functionality.

If the unsuspecting user places the macro virus on their home computer via diskette, many problems can occur. Not only can the home computer become infected, but the network could also be re-infected. After modifying the infected file at home, the user can bring the file back to the office and infect his individual workstation. However, since the virus definitions should have been updated, the malicious code protection at the workstation should identify the virus and remove it. The user should then scan the home computer and remove all infections on that computer as well.

## 6.6.8.2 Case 2: Polymorphic Virus Attack

Polymorphic viruses increasingly represent serious threats to computer networks. Prevention, detection, containment, and recovery from potentially lethal polymorphic computer viruses should be an important task of every user, network administrator, and senior management officer. Establishment of an adhered to anti-virus computer policy is a must for everyone requiring any degree of protection for their systems against polymorphic virus attacks.

To successfully prevent polymorphic viruses from entering into a computer system, potential vulnerabilities must be identified and eliminated. Attackers often look to exploit the most obvious vulnerability of a computer network. Inadequate security mechanisms allow unauthorized users entry into computer systems, potentially allowing data to be compromised, replaced, or destroyed. Determent of attackers can be accomplished by having a predetermined computer protection plan in place. Also, contingency plans will enable the containment and eventual recovery from a polymorphic virus attack. Another technique for preventing polymorphic virus attacks is to set up false data directories or repositories to fool the attacker. (See Section 6.6.5.1 Types of Malicious Code, Polymorphic Viruses.) Preparation for any incident of an attack and knowledge of how a given attack might occur is all part of the strategic virus protection plan that should be implemented prior to operation of a computer network.

Detection of polymorphic viruses becomes exponentially easier when the polymorphic virus signature is cataloged in an anti-virus definition table and updated regularly to all systems gateways. This can happen in one of two ways. A user can notice altered functionality on a workstation, and after technicians investigate the problem, the polymorphic virus is finally discovered. Then, technicians inform vendors who update the virus definitions for others. A user can also remove the polymorphic virus after vendors have updated their virus definitions by downloading the newest virus definitions and scanning the entire system. Establishment of an updating policy not only for system gateways, but also for individual computer workstations, greatly increases the likelihood of preventing a polymorphic virus from entering and replicating itself on a given network.

Recovery methodologies are integral to the overall readiness of an anti-virus prevention plan. Even the best prepared plans sometimes fail. Having written procedures in place to recover from a catastrophic event could mean the difference between a company surviving or going out of business. Recovery consists of virus-free tape backups of recent data, providing an environment free from all viruses, and establishing the network to pre-virus infection operation. There are inexpensive software applications that unobtrusively tracks disk activity in such a way that it can return a system to precisely the way it was prior to a computer virus incident. Backing up data or implementation of a mirroring solution is key to having a ready alternative source of providing information to users on a moments notice. Unless uniformly adopted throughout the entire organization, a plan will have little chance of ever becoming successful. Dedicated personnel responsible for predetermined actions in anticipated situations are crucial for the protection of computer systems.

### 6.6.8.3 Case 3: Trojan Horse Attack

Eradication of a Trojan Horse encompasses many of the same procedures taken to eradicate macro and polymorphic viruses. (See sections 6.6.8.1 Case 1: Macro Virus Attack & 6.6.8.2 Case 2: Polymorphic Virus Attack.). This is because the Trojan Horse can contain a virus inside of the apparently harmless program. However, in this case, something else must be done to rid the network of the sniffer program hidden inside the Trojan Horse. There is no one solution to prevent, detect or remove sniffers. Since sniffer programs are extremely difficult to detect, the first level of defense against them is to make sniffing difficult. The network should use a switch instead of a hub to protect sniffing of internal user passwords. By using an encryption mechanism for message transmissions and e-mail transactions, sniffing of important data such

as passwords can be prevented. The use of "ssh" or other encrypted commands can help keep passwords private. Another precaution against password sniffing in the use of 1 time passwords. It does an attacker no benefit to sniff a password that is only valid during a very short time period.

In this case, the presence of sniffers is suspected since numerous forged e-mails have occurred. By applying the above measures of encryption and secure commands, sniffers can be rendered ineffective as passwords become much harder to decipher. It is also a good practice to change passwords often, or have the system administrator force users to change their passwords periodically to decrease the chance sniffer program users have time to decrypt encrypted passwords.

Also, it can not be stressed enough how important it is to establish a complete and comprehensive malicious code protection backup system. If sniffer program users gain unauthorized access to the network, user applications and data files could be deleted. The only countermeasure in this case is to change all passwords and restore the system to prior functionality from full system backups. However, when systems are restored the sniffer must not be restored also.

## References

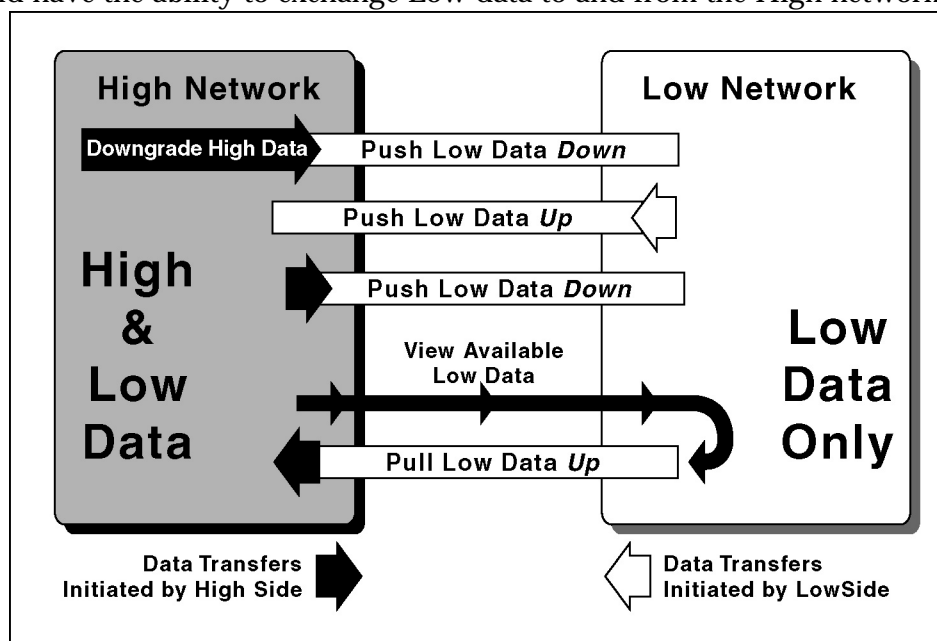
- “A Clear and Present Danger,” Information Week. May 22, 2000, p.166.
- “AINT Misbehaving: a Taxonomy of Anti-Intrusion Techniques,” SANS Institute Resources Intrusion Detection FAQ. Ver. 1.33.
- Bassham, Lawrence E. & Polk W. Timothy, “Threat Assessment of Malicious Code and Human Computer Threats,” NIST – Computer Security Division, October 1992.
- “Batten Down The Digital Hatches!” Forbes. June 12, 2000 p.246.
- CIAC, “H-05 Internet Hoaxes: PKZ300, Irina, goot Times, Deeyenda, Ghost,” U.S. Department of Energy, Nov 20, 1996.
- Chess, David., “The Future of Viruses on the Internet,” Virus Bulletin International Conference In San Francisco, October 1997.
- “DANGEROUS 'LOVE': Recent virus attacks prompt enhanced security measures,” Computer Reseller News. May 29, 2000, p.45.
- “Don't fall for a Virus Hoax,” Sophos Virus Info, 23 Nov. 1999.
- F-Secure, “Security Risks for the Road Warrior,” Wed. July 12, 2000.
- “Frost & Sullivan Awards Internet Security Systems the 2000 Market Engineering Marketing Strategy Award,” Press Release. June 28, 2000.
- Gabrielson, Bruce C., “Computer Viruses,” INFOSEC Engineering, AFCEA Seminar, Burke, VA. Sept. 1994.
- “An Introduction to Computer Viruses (and other Destructive Programs),” McAfee Network Security and Management.
- Ludwig, Mark., The Giant Black Book of Computer Viruses, Show Low, AZ, 1995.
- McAfee, John., Computer Viruses, Worms, Data Diddlers, Killer Programs, and Other Threats to Your System, Fifth Avenue, NY, 1989.
- Micro, Trend., “Eliminating Viruses in the Lotus Notes Environment,” 1999.
- “Securing dot-com – New viruses, distributed security threats pose perpetual challenges to IT,” eWeek, June 26, 2000 p.1.
- Slade, Robert M., “Antiviral Protection Comparison Reviews,” 1995.
- Wack, John P. & Carnahan, Lisa J., “Computer Viruses and Related Threats: A Management Guide,” NIST Special Publication.
- “Understanding Symantec's Anti-virus Strategy for Internet Gateways,” The Symantec Enterprise Papers, Volume XXX.
- “Understanding and Managing Polymorphic Viruses,” The Symantic Enterprise Papers, Volume XXX.
- “What Virus Is Lurking?—Better not touch that E-mail.” Computer Reseller News. June 5, 2000 p.1.

## 6.7 Multi-Level Security (MLS)

### 6.7.1 High-to-Low

The High-to-Low category is a subcategory of MLS. The goal of this category is to provide solutions giving installations the ability to connect networks of unlike classification (in generic terms, the classifications can be described as “High” and “Low”), as depicted in Figure 6.7-1. Given that the classifications of the data on the two networks is ordered, i.e., one is higher than the other is, users would have the ability to exchange Low data to and from the High network.

This ability is in spite of the fact that neither the High network nor the Low network has the ability to label the data. All data on the High side is considered to be High data. Users on the High network must explicitly designate data as Low and then request that it be transferred to the Low network. This is a flow of Low data from High to Low. Likewise, Low data may flow from Low to High as a result of a user on the



**Figure 6.7-1 High To Low Concepts**

Low network sending data to the High network (e.g., in an electronic mail (e-mail) message), or a user on the High network requesting data from the Low network, e.g., through a HyperText Transfer Protocol (HTTP) request to a web server on the Low side.

In no case is it desired for High data to cross between the two networks in either direction. There are three primary statements within the policy for High-to-Low. First, the High data on the High network must never cross to the Low network. Second, the High network must be protected from attacks that could cause High data to be leaked to, modified by, or destroyed by users on the Low network. Third, High network resources may not be utilized, modified, destroyed, or made unavailable by unauthorized Low network users.

These requirements apply to all High-to-Low connections, regardless of the actual classifications. Possible scenarios include Secret-to-Unclassified, Secret United States (U.S.)-to-Secret Allied, Top Secret-to-Secret, as well as those that are not formally classified such as Sensitive But Unclassified (SBU)-to-Unclassified Internet. It is the intention of this Framework to specify

requirements in a form that is generic enough to address all popular network services, e.g., E-mail, HTTP, File Transfer Protocol (FTP), Database. The requirements will be phrased in terms of “pushing” and “pulling” data between the two networks.

### 6.7.1.1 Target Environment

There are three target environments that this Framework will address:

- 1) Allow users on the High network to push Low data to users on the Low network, and allow users on the Low network to push Low data to users on the High network.
- 2) Allow users on the High network to downgrade data to Low, and push that data to a server on the Low network for subsequent pull by users on the Low network.
- 3) Allow users on the High network to view and import (pull) data that exists on the Low network.

In the remainder of this Framework, the above three capabilities will be referred to as:

- Communication,
- Releasability, and
- Network access

respectively.

### 6.7.1.2 Consolidated Functional Requirements

#### 6.7.1.2.1 Requirements for “Communication”

Current requirements are:

- Send and receive electronic mail between the High network and the Low network.
- E-mail must conform to standards used in the wider community.
- E-mail must allow users to send and receive attachments in both directions.

Anticipated requirements are:

- Enable users to use Chat as a means of communication between High and Low network users.
- Enable Internet telephony between High network users and Low network users as the technology becomes available.
- Enable video conferencing between High network users and the Low network users.

#### 6.7.1.2.2 Requirements for Releasability

Current requirements are:

- Enable authorized users on the High network to designate and push—e.g. FTP, E-mail, HTTP Post, etc.—data to the Low network that is releasable to users on the Low network.

- Enable authorized users on the Low network to access the released data using Web technology, FTP, data base access techniques.
- Released data may be restricted to certain users, or it may be made publicly available.
- Released data may be text, video, images, audio, or executable software.

### 6.7.1.2.3 Requirements for Access

Current requirements are:

- Users on the High network must be able to access the vast information resources on the Low network.
- Access methods may be HTTP, FTP, Gopher, Wide Area Information Service (WAIS), SQL, or Web Push. With Web push, as a result of a previous High to Low access request, information is pushed onto the High network from the Low network.

### 6.7.1.3 Attacks and Potential Countermeasures

The following section itemizes previously identified attacks that were explained in Chapter 3 (System Security Methodology) of this document, and matches these attacks with potential countermeasures that may be included in solutions addressing the High-to-Low requirement category.

#### 6.7.1.3.1 Passive Attacks

- *Traffic Analysis* – As of now, no technical countermeasure has been identified that is appropriate for inclusion in High-to-Low requirement category solutions.
- *Monitoring Plaintext* – The appropriate countermeasure to this attack is to deny access to the data by unauthorized users by encrypting the data or by using other data separation techniques that will restrict unauthorized release of data. (Note that utilizing encryption is possible only when both parties have access to the same algorithms and keys and the same capability to encrypt and decrypt the data properly.)
- *Decrypting Weakly Encrypted Traffic* – Countermeasures are to use adequate encryption algorithms, and maintain sound key management.

#### 6.7.1.3.2 Network-Based Attacks

- *Modification of Data in Transit* – The countermeasure to this attack is to use digital signatures or keyed hash integrity checks to detect unauthorized modification to the data in transit.
- *Insertion of Data* – There are many countermeasures to the malicious insertion of data. They include the use of timestamps and sequence numbers, along with cryptographic binding of data to a user identity, to prevent replay of previously transmitted legitimate data. Data separation or partitioning techniques such as those used by firewalls and guards denies or restricts direct access and the ability to insert data by Low side agents into the High side network.

- *Insertion of Code* – Virus scanning by High side users and enclave protection devices attempts to detect incoming viruses. Cryptographically authenticated access controls may be utilized to allow data only from authorized sources to enter the High network. Audit and intrusion detection techniques may detect breaches in established security policy and anomalies.
- *Defeating Login Mechanisms* – The most appropriate countermeasure for this is cryptographic authentication of session establishment requests.
- *Session Hijacking* – The countermeasure for this is continuous authentication through digital signatures affixed to packets, or at the application layer, or both.
- *Establishment of Unauthorized Network Connections* – There is no technical countermeasure for this. It is incumbent on the management and administration of the local network to prohibit unauthorized connections between High and Low networks, and to enforce that policy through non-technical means. Various commercial tools may be utilized by system administrator personnel to detect such connections.
- *Masquerading as an Authorized User* – The appropriate countermeasure is to use cryptographic authentication in conjunction with timestamps or sequence numbers to prevent replay of authentication data. Another countermeasure to prevent stealing an authentic session is to cryptographically bind authentication data to the entire session/transaction.
- *Manipulation of Data on the High Side* – The appropriate countermeasure is to permit only authorized users to access the data on the High side using cryptographic authentication and data separation techniques.

### 6.7.1.3.3 Insider Attacks

- *Modification of Data or Modification of Security Mechanisms by Insiders* – The primary technical countermeasure is to implement auditing of all security relevant actions taken by users. Auditing must be supported by timely, diligent review and analysis of the audit logs generated. Other countermeasures to these attacks are non-technical and therefore not addressed by the High-to-Low requirement category solutions. Non-technical countermeasures include personnel security and physical procedures.
- *Physical Theft of Data* – Again, the countermeasures to these attacks are non-technical and therefore not addressed by the High-to-Low requirement category solutions. Appropriate non-technical countermeasures include personnel security and physical security procedures, which inhibit actual removal of data, either in printed form or on storage media.
- *Covert Channels* – The countermeasure against a covert channel between the high and low networks, is a trusted guard function that examines network header fields and network messages for possible unauthorized information.

### 6.7.1.3.4 Development and Production/Distribution Attacks

- *Modification of Software During Development, Prior to Production* – The countermeasures for threats during this phase include use of strong development processes/criteria such as Trusted Software Development Methodology and subsequent evaluation of software



by third party testing using high assurance methods and criteria such as the Trusted Product Evaluation Program (TPEP) and Common Criteria testing.

- *Malicious Software Modification During Production and/or Distribution* – The countermeasures for threats during this phase include high assurance configuration control, cryptographic signatures over tested software products, use of tamper detection technologies during packaging, use of authorized couriers and approved carriers, and use of blind-buy techniques.

### 6.7.1.4 Technology Assessment

This section discusses general technology areas that can be used in system solutions to address the functional and related security requirements associated with the High-to-Low requirement category. Section 6.3.1.5 (Requirement Cases) proposes various system level solutions that build upon these general technology areas. The proposed security countermeasures included in each system solution result from our analysis of user target environments, functional requirements applicable to the *communications*, *releasability* and *network access* requirements, and attacks and potential countermeasures as have been discussed in previous sections.

The framework divides the technology of protection between High and Low networks into three categories:

- 1) Data Separation Technologies,
- 2) Authenticated Parties Technologies, and
- 3) Data Processing, Filtering and Blocking Technologies.

This categorization allows us to make some high level assessment of system assurance provided for groups of similar solutions, thereby ordering solutions in terms of security robustness. These three generic categories of potential solutions are explained in more detail in subsequent paragraphs of this section.

#### 6.7.1.4.1 Data Separation Technologies

System solutions that would logically fit into this technology category would allow users who are located in High side protected enclave environments to have access to both High network and Low network data, but prohibit pushing and pulling of data between these two networks. Typically, solutions in this category rely upon physical separation of data (from user interface to redundant distribution networks) in order to provide data segregation between High and Low applications.

In most cases High-side users are restricted from using sophisticated automated means that allow for the storage or manipulation of Low-side generated data on the High network. In addition, High-side users are also restricted from directly extracting Low data from the High network applications, or using a broad range of applications to move the extracted data to the Low network.

All of the proposed solutions that are included in this category do however provide for the data transfer techniques previously described as *communications*, *releasability* and *network access*, but do so only within networks of the same level.

For *communications* exchanges, typical solutions in this category allow access for High side users to redundant network access points, which are individually connected to both networks, i.e., High network users have access to two network access points, one for the High network and one for the Low network. Users may have two processors with shared monitors and keyboards or several users may be provided access to a shared Low network interface located in a centralized location. Likewise for both *releasability* and *network access* exchanges, users on the High network side will interface to logically separated network interfaces.

The economics of solutions that fit into this category must be examined and a tradeoff analysis completed which compares the savings resulting from greatly simplified security mechanisms and reduced complexity of security management infrastructure and personnel support, to the cost of redundant local networks and network management. The primary advantage of data separation solutions is that all of the solutions in this category provide the highest degree of system level security, and may in fact be the only solutions that are acceptable for very high assurance networking requirements. These are very secure system topologies, providing the best protection from both passive and network attacks.

These solutions do not allow data to flow between the High network and the Low network. Hence, they are robust in preventing attack of the High network and leakage of High data to the Low network. The only true data separation technology is physical isolation of the network. Any connection between the two networks will create the potential of at least minimal leakage via covert channels, as well as the operational risk of attacks from Low to High. Solutions here include:

- Isolated Networks,
- Secure Network Computers,
- Starlight Interactive Link, and
- Compartmented Mode Workstation (CMW).

Each of these is discussed below.

## Isolated Networks

This solution is simply to maintain two networks, one for High data and one for Low data. The two networks are never to be connected together. This would require redundant infrastructures, at additional cost. However, the cost can be justified in environments where users cannot tolerate the risk that the High data might be compromised or the High network attacked.

The number of workstations on each network is a function of the need within the organization to have individuals with access to both networks. Perhaps the Low network can be accessed via shared workstations if it is not necessary for all users to have access from their desktops.

The specific capabilities addressed by this solution are communication and network access. Automated releasability to the Low network of data created on the High network is not addressed by this technique. Regrading and subsequent release to a co-located Low network computer, of information contained on the High network computer may be performed by overt human intervention, e.g., human review and retyping of data on the Low network computer or optical scanning. Communication and network access are addressed by allowing the user who has access to a terminal for each network to exchange electronic mail, participate in Chat ses-

sions, and perform World Wide Web (WWW) browsing with other parties on either network by using the appropriate terminal.

While many customers wish to avoid using separate networks, this option bears consideration with the increased availability of low cost Personal Computers (PCs) and network computers. The cost of implementing and operating two separate networks might actually be less than implementing and managing sophisticated network security systems. Furthermore, the richness of the network access will be unimpaired by the security at the boundary of the High network.

## Secure Network Computers

Research is being done on a secure network computer that will employ a cryptographic token to separate data on the network. The concept is that the network will be classified for Low data, while having servers connected that process High data. All High data on the network is encrypted to provide separation. The workstations on the network are all *single level at a time* with only volatile memory. They are network computers that accept a cryptographic token to encrypt and decrypt all communications over the network. Depending on the token placed in a network computer at any time, it will be able to access either High servers or Low servers, but not both. When the token is changed, the volatile memory of the network computer is cleared. Since this is a research project, no commercial products are yet available. Hence, this is identified as a technology gap that is being worked.

When secure network computers become available, they will allow communication and network access on High networks and Low networks using the same device. They will not allow automated regrading of data, so it would not be possible to forward an e-mail message from the Low network to recipients on the High network. Likewise, the secure network computer does not support automated releasing of Low data from the High network. To release Low data residing on the High network, users would be required to perform a human regrade procedure, using non-automated methods such as re-typing of the data or optical scanning.

## Starlight Interactive Link

This is a technology that is being developed in Australia that allows a single monitor, mouse and keyboard to have access to two different computers. One computer is connected to the High network, and one is connected to the Low network. The technology allows *single level at a time* access to the two networks from a single location. Data does not transfer between the two without human review. It is possible to cut-and-paste data from Low to High only (never High to Low) using the standard X Windows cut and paste capability. This can be done only with human intervention. There is no way to automate the regrading of data. It should be noted that the cut-and-paste Low to High capability introduces risk that the data pasted to the High network could contain malicious code.

The implementation employs a one-way fiber optic link with the Low computer. This prohibits data leakage from High to Low. Because of the fiber optic link, data can only flow away from the Low computer to the display; it can never flow from the display to the Low computer.

The Starlight Interactive Link supports communication and network access from a single location. It does not support automated releasability from the High network to the Low network.

Since the Starlight Interactive Link is not yet a commercial product, it is identified as a technology gap.

## Compartmented Mode Workstations (CMW)

Another solution in the data separation class is to use CMWs or higher assurance workstations, if available. These could be judiciously allocated to the users who need to access both the High network and the Low network. With this approach, each user is then able to access both the High network and the Low network.

The specific capabilities addressed by this solution are communication, network access, and releasability. Communication and network access are addressed by allowing the user who has access to a CMW, which is connected to each network, to exchange electronic mail, participate in Chat sessions, perform WWW browsing with other parties on either network by using a window dedicated to the proper network. Releasability and communication between the High network and the Low network are addressed by the CMW *cut-and-paste* and *downgrade capability*. This operation allows users to highlight information in a High window and use the cut or copy command to place it in a buffer for review. The resulting information is then downgraded, appropriately classification marked, and displayed to the user in a Low window for visual review and release.

Cut and paste between sensitivity levels is an action that requires the CMW to be configured with this privilege; it is not allowed by default. If the CMW is not configured with this privilege, complete logical data separation is achieved.

### 6.7.1.4.2 Authenticated Parties Technologies

System solutions that would logically fit within this category are solutions that mandate the use of cryptographic authentication mechanisms prior to allowing access. Examples of actions that could be governed by this technology are:

- Allowing High users to access servers on the Low network when the servers can be authenticated.
- Allowing High users to release data from the High network based on their authenticated identity.
- Allowing Low data to enter the High network when the Low data is cryptographically bound to an authorized individual through a digital signature.

Authenticated access is widely available and is supported by a large number of standards and protocols. It allows two parties that intend to exchange data to identify themselves to one another and positively authenticate their identities. Hence, they become mutual trusting parties. The data that flows between these trusting parties is at the level of the lower party. This paradigm is applicable to the previously discussed modes of data exchange: *communication*, *releasability*, and *network access*.

Authenticated access solutions typically address *communication* data exchanges by use of digital signatures for electronic mail messaging applications, e.g., Message Security Protocol (MSP) or Secure/Multipurpose Internet Mail Extension (S/MIME). Such solutions typically involve the concept of protected enclaves for the system-high users that are separated from the system-low network users by some sort of enclave boundary protection device such as a guard or firewall.

In such a topology, Low network users might utilize digital signature technology to authenticate themselves to High network users. Also, the guard might incorporate Access Control List (ACL) mechanisms to make access decisions governing the set of users that are authorized to release information from the High network. Access control lists can also be used to restrict the set of Low network users that are authorized to push data up to the High network.

Likewise, authentication solutions are applicable to *releasability* data exchanges in that the releaser can digitally sign data to be released. Again, enclave boundary protection systems such as guards might utilize ACLs that would regulate who in the system-high network is authorized to release data from the High side network. The enclave boundary protection system might also perform content review of the data submitted for release.

Lastly, authentication solutions are applicable to *network access* data exchanges typically through the use of Commercial-Off-The-Shelf (COTS) protocols such as Secure Sockets Layer (SSL), Secure HyperText Transfer Protocol (S-HTTP), SOCKS, Secure Electronic Transaction (SET), and Internet Protocol Security (IPSec) for Web access, database access, FTP access, etc.

It is logical to conclude that security is enhanced if parties that are mutually trusting create a closed virtual community. The down side of these types of solutions is that, in general, they mandate that both parties have compatible security mechanisms to strongly authenticate themselves to one another. Therefore, the implication is that the number of Low network resources that are accessible is greatly reduced to include only those that are “security enabled.” In the case of *network access* requirements, the requirement to be security enabled may greatly reduce the availability of access to public information resources.

It must also be noted that authentication solution topologies normally necessitate a very restrictive policy whereby activity is allowed only with other parties that are authenticated as part of the closed, and therefore trusted, community. Conversely, if the community is opened by a single party who interacts with another party outside of that community, then the entire community is potentially vulnerable to attack.

While authentication technologies are widely available, they have yet to become fully mature. For a discussion of hurdles that must be overcome, see Section 6.3.1.4 (Technology Gaps).

Solutions using Authenticated Parties include the following.

- Authentication between clients and servers using SSL.
- Host to host authentication using IPSEC with the Authentication Header.
- Authentication at the application layer.

These are discussed below.

## Authentication between Clients and Servers Using SSL

SSL[1] is becoming a popular security protocol for implementing privacy and authentication between communicating applications. It is a transport layer security protocol, enabling the encryption and authentication of arbitrary applications. The protocol prevents eavesdropping, tampering with information, and forging of information sent over the Internet.

The SSL protocol includes a lower level protocol (called the SSL Record Protocol) that encapsulates higher-level security protocols. The SSL Handshake Protocol is one such encapsulated

protocol. It allows communicating parties to authenticate one another, and to establish cryptographic algorithms and keys at the start of a communication session.

Connections using SSL have three properties:

- The communication is private. The initial handshake uses public key cryptography to define a secret key. The secret key is then used with symmetric cryptography to encrypt all communications.
- Clients and servers can authenticate one another during the handshake using public key cryptography.
- The entire communication is protected against tampering or insertion of data. Each datagram has a Message Authentication Code that is a keyed hash value.

The SSL protocol can be used for network access between clients on the High side and servers on the Low side. This can give confidence that the server is trusted to some degree. A policy requiring that SSL be used for all network access between High and Low would effectively permit access only to servers on the Low side that have the ability to authenticate using SSL. However, such a policy might not be useful if there are some Low servers that have the ability to authenticate, but should not be included within the set of servers to which access is allowed. The goal should not only be authentication. Rather, the goal should be access control, with authentication being a means to implement access control. This is accomplished by maintaining a list of Low servers that, once authenticated, can be accessed by High clients. That list is best maintained by an enclave boundary protection system, e.g., guards.

If an enclave boundary protection system is in use, SSL can be used between the enclave boundary and the Low server. If the SSL is between an enclave boundary protection system and the Low server, then guarding, filtering, and blocking technologies can also be applied to allow access to only those Low servers that are on an access control list. The enclave boundary protection system would keep a list of servers to which network access is allowed, and would enforce the policy that no network access is allowed to any other servers. SSL could also be used as a basis for communication via e-mail, Chat, Whiteboarding, or other protocols, since it is a transport layer protocol and is independent of the application. Since SSL also gives the capability to encrypt all application layer data, the communication between the enclave boundary and the Low server is private.

SSL can also be used between the client on the High network and the enclave boundary. This allows the enclave boundary protection system to maintain a list of High clients that are authorized to communicate with users on the Low network, to access information on the Low network, and to release information to the Low network.

Using SSL for end-to-end encryption and authentication from High clients to Low servers limits the effectiveness of an enclave boundary protection system. In this case, the enclave boundary protection system cannot see the application layer information being communicated between the client and the server. Therefore it can make access control decisions only on information in the transport layer and layers lower than the transport layer. Thus, a tradeoff must be made between end-to-end security and the access control capabilities of an enclave boundary protection system. However, the benefits of using an enclave boundary system to enforce access control can be argued to outweigh the loss of uninterrupted end-to-end encryption and authentication.

For High to Low, the optimal use of SSL is to have two SSL connections meeting at the enclave boundary protection system. One connection is between the High host and the enclave boundary; another is between the enclave boundary and the Low host. This allows the enclave boundary protection system to perform filtering, authentication, access control, and auditing of all traffic passing from High to Low. To perform this function, the enclave boundary system would use a proxy that effectively glues two separate SSL sessions together.

## Host-to-Host Authentication Using IPSEC with the Authentication Header

Like SSL, the IPSEC security protocols allow encryption and authentication of all information above the network layer in the Transmission Control Protocol (TCP)/IP stack. Unlike SSL, Internet Protocol Security (IPSec) resides at a lower layer in the communication stack, and has the capability to completely encapsulate IP packets, including the source and destination addresses. Where SSL can be described as a process-to-process security protocol, IPSec is sometimes referred to as a host-to-host security protocol.

In connections between High networks and Low networks, IPSec can be useful in authenticating the hosts at the communication endpoint, and in giving privacy of the data being transmitted. Since IPSec is at a lower layer in the communication stack than SSL, IPSec can help in prevention of spoofed IP addresses.

IPSEC is of little use in High to Low connections without an enclave boundary protection system at the point where the High network is connected to the Low network. The enclave boundary protection system is needed to perform access control between High and Low. At the same time, the enclave boundary protection system is rendered useless if IPSec with encryption is used between the High host and the Low host, since the communications would be encrypted with a key private to those two endpoints. For High to Low, the best use of IPSEC is between the Low host and the enclave boundary protection system, and also between the High host and the enclave boundary protection system. This allows the enclave boundary protection system to authenticate both endpoints of the communication, although it creates a complexity in key management for the enclave boundary protection system. Since most enclave boundary protection systems that are suitable for High to Low do not perform IPSec, this is considered a technology gap.

## Authentication at the Application Layer via Digital Signatures

Current High to Low solutions for electronic mail have the capability for digital signatures to identify the originator of e-mail messages. These solutions also depend heavily on a mail guard for enclave boundary protection. Like SSL and IPSec, the enclave boundary protection system cannot perform the functions of inspecting the content of the message or verifying the digital signature if the message is encrypted. The currently available e-mail solutions allow the guard to decrypt a copy of outgoing messages in order to perform filtering on the contents of those messages.

Authentication at the application layer using digital signatures allows the enclave boundary protection system to determine the individual that is responsible for the traffic passing from High to Low, and then make an access control decision to allow or disallow the traffic. Since

the digital signature is based on public key cryptography, a public key infrastructure must be in place to enable this solution.

### 6.7.1.4.3 Processing, Filtering, and Blocking Technologies

Solutions that logically fit within this solution category utilize various processing, filtering, and data blocking techniques in an attempt to provide data sanitization or separation between High network data/users and Low network data/users. Data originating from the High network is assumed to be High data though it may be asserted to be Low data by a High network user. Automated processing and filtering techniques may be performed by enclave boundary protection devices such as a guard, and if such tests are successfully passed, the data is actually regraded by automated means. In the reverse direction, such solutions often incorporate data blocking techniques, (typically in firewalls but also in guards) to regulate the transfer of data from Low network users to High network users. Use of certain protocols may be blocked and/or data may be processed or filtered in an attempt to eliminate or identify viruses and other malicious code transfers.

The technology categories of data separation and authenticated parties do not allow users to use automated means to transfer data between the High and the Low network. The only technology that allows automated data regrading and transfer is processing, filtering, and blocking. Hence, this technology is the linchpin of High to Low. Without processing, filtering, and blocking techniques, there are no automated mechanisms supporting the regrading of information from High networks to Low networks. Data separation and authenticated parties technologies are restricted to allowing information transfer between networks only by means of human intervention such as retyping or optical scanning.

It must be emphasized that data transfer between High and Low involves risk, and one must take steps to mitigate risk. If data separation via a technology described in any of the other solution categories is not possible, then processing, filtering, and blocking must be considered. It must however be recognized by implementing organizations that these techniques involve inexact attempts to filter High data from outgoing transmission through content checking against a pre-defined list of prohibited strings. It also involves scanning for and detecting virus-infected executables, and blocking executables. Since there are an almost infinite number of possible executables, and malicious ones can be detected only through prior knowledge of their existence, the problem of detecting "maliciousness" in an arbitrary executable is not computable. This is exacerbated by the fact that there exist many executables that users wish to allow to cross the network boundary (e.g., Java applets, Active X controls, JavaScript, Word macros) and that they would therefore not wish to filter out or block. Only by performing a detailed risk management tradeoff analysis wherein operational needs are weighed against security concerns can these issues be resolved.

Solutions using processing, filtering, and blocking employ some type of processing to allow information flow between the two networks, but attempt to detect and block attacks and High data leakage. Solutions here include:

- I-Server for Communication, Network Access and Releasability,
- Mail Guard, and
- Low-to-High Replication.



Each of these is discussed below.

## I-Server for Communication, Network Access, and Releasability

This solution uses a special purpose computer, dual-homed at the boundary between the High network and the Low network. The solution is identified as a technology gap due to the non-existence of commercial products that have this capability. The technology needed to develop such products is well understood, however. The computer, called an *Intermediate Server*, is a remote host that users on the High network can login to and execute browsers and Internet client software. The *I-server* is ideally a trusted computer with the ability to keep data of differing classifications separated. It also has the ability to protect itself against attack from the outside. Malicious code that might execute as part of Java applets or Active X controls would not be able to damage the I-server or the High network due to rigid design constraints.

The I-server is protected by a robust architecture that prevents tampering or modification of the operating system. This architecture also constrains the processes that are running any hostile executables to their own address space, and gives them no privileges to observe or modify files. The High network is protected by the remote location of the I-server, keeping potentially hostile code off of the High workstations and servers. Only the display of the information retrieved from the Low network is sent to the High network.

The specific capabilities addressed by this solution are communication, network access, and releasability. Communication is addressed by allowing the user on the High network to exchange electronic mail with users on the Low network, and to participate in Chat sessions with parties on the Low network. Network access is addressed by allowing users on the High network to perform WWW browsing via the I-server, and to access FTP servers on the Low network via the I-server. Releasability is addressed by allowing users on the High network to upload files to be released to the I-server, applying filters to determine that the information is indeed releasable, and then sending the released files to external servers.

The I-server architecture enables indirect accesses to the Low network. The I-server is a trusted computer that has MLS capability with high assurance. The I-server is connected both to the Low network and to the High network. Users on the High network log on to the I-server at the Low level. Browsers and other Internet clients, e.g. Simple Mail Transfer Protocol (SMTP), FTP, and Telnet; execute on the I-server, and all information retrieved from the Low network stays on the I-server at the Low level. That information can be viewed by the user on the High network who requested it. The viewing is done through a terminal emulation protocol between the I-server and the user workstation on the High network. Since the I-server is a trusted computer that can protect itself from attack, the threat posed by malicious executables is greatly diminished.

The following are the steps a user would perform to browse the Low network from the High network through an I-server.

- Login to the I-server at the Low level.
- Authenticate to the I-server via password or other authentication mechanism.
- Run the Web client available on the I-server.
- Type in the URL/IP address desired or select from your personal set of bookmarks/favorites or select entries from an address book.

- See the responses through terminal emulation at the user's workstation, and, if desired, save them on the I-server for future reference. Files saved on the I-server will be saved at the Low level.

Note that the steps above do not include a means for a user to pull data retrieved from the Low network to his/her workstation on the High network. Since pulling of data from the Low network could impose an avenue for attack, the I-server prohibits this pulling. To allow this pulling of information through the I-server would bring along the inherent risks of pulling data from untrusted sources on the Low network. If pulling of data is a user requirement, then procedures and policies must be in place to mitigate risk of pulling hostile executables. One such policy would be to allow pulling of only ASCII text, and prohibit use of decoding software (such as UUdecode) on that text.

The main security weakness of the I-server is the potential for leakage of data from the workstation on the High network that is untrusted, to the Low process executing on behalf of the user on the I-server. This could occur through a covert channel in the terminal emulation protocol, and be driven by a Trojan Horse on the user's workstation. It would also require collusion at the receiving end (the Low process on the I-server). This vulnerability would be difficult to exploit, and therefore is considered lower risk than would be present if the HTTP protocol were being sent end-to-end between the workstation on the High network and the server on the Low network.

## Mail Guard

This solution is readily available with both commercial and government-developed products. The guard is deployed at the boundary of the High network and the Low network. The guard performs filtering and control of mail messages passing High to Low and Low to High. The filtering is based on the headers of the mail messages, e.g., sender, recipient, presence of signature; as well as the contents of the mail message, e.g., encryption of contents, presence of prohibited words or phrases. At this time the solution only addresses communication via electronic mail. Guards are typically used in conjunction with "authenticated parties" technology. This adds some strength to the relative weakness of content filtering employed by a guard.

Current mail guards are very flexible, allowing implementation of a wide variety of message acceptance and message release policies. It is possible to configure mail guards to be very liberal in these policies. Policy makers must pay strict attention to policy decisions to assure that policies are not so liberal as to negate the usefulness of the mail guard.

## Low-to-High Replication

Low-to-High Replication<sup>1</sup> allows users on the High network to receive data that originates on the Low network, without having to explicitly request that the data be sent from the Low servers. Replication can be used for network access, pushing data from the Low network to the High network. It cannot be used for releasability or for communication, because its primary security property is the prevention of data flows from High to Low.

Replication can give the High network any application that passes messages from one host to another. Examples are database replication, FTP, electronic mail, and Web Push protocols.

To prevent data leakage from High to Low, replication does not allow a direct back channel to send message acknowledgements from the High network to the Low network. To do so would

allow quite a large covert channel. The replication acts as an intermediary, sending acknowledgements to the Low sender, and receiving acknowledgements from the High recipient. The Low sender cannot determine with precision the timing of the acknowledgements sent from the High side. Hence, the bandwidth of the back channel is reduced by the intermediate buffer within the replication process. This disconnects any direct communication from High to Low.

Replication does not mitigate the potential risk that data replicated into the High network might be hostile executable code. Mitigation of this risk would require that data be replicated first in a network guard that inspects the data for potentially hostile code, making sure the data passes this inspection before being forwarded into the High network.

### 6.7.1.5 Requirements Cases

This section is intended to address the connection of High to Low networks for purposes of communication, network access and releasability. These are general, functional requirements that have been articulated by various customers. Presently, only the Secret-to-Unclassified network connection scenario has been analyzed in detail. There are other connection scenarios where similar requirements appear to be appropriate. The additional scenarios we are aware of are Top Secret-to-Compartmented-Top Secret, Top Secret-to-Secret, and Secret U.S.-to-Secret (Allied). These other scenarios are under analysis, and their requirements will be presented in future versions of the framework if they are found to be different from the Secret-Unclassified case.

#### Case 1: Secret-to-Unclassified

Users on the Secret network have a need to connect to the Unclassified network for the purposes of communication, network access and releasability. For communication, the needed application is electronic mail. Access to the Unclassified network is needed also via Web protocols, using commercially available web browsers. Finally, Secret users sometimes create large files that are in reality Unclassified. In some cases users have a need to release these Unclassified files to the Unclassified network.

Electronic mail is currently enabled between Secret and Unclassified in many instances through a mail guard, which is sometimes coupled with a Commercial Off-The-Shelf (COTS) firewall. In the Defense Message System, electronic mail will be enabled between Secret and Unclassified using a mail guard. The immediate need is to develop the additional capability to use Web-based protocols (i.e., HTTP) to access web servers on the Unclassified network. Another immediate need is to develop the capability to release large files from Secret to Unclassified (probably using FTP). Current guards do not have the capability to allow network access and releasability.

The environmental requirements for the Secret-to-Unclassified connection include:

- Secret users must be able to use COTS software, e.g., browsers and e-mail clients, in accessing information, communicating with users, and releasing information on the Unclassified network.
- Secret users must be able to use the installed base of operating systems, whether they are Windows or Unix.

- The new capabilities for access to the Unclassified network and for releasability must coexist with existing capabilities to send and receive e-mail with users on the Unclassified network.

## Case 2: Secret US-to-Secret Allied

This section will be provided in a later release of the Framework.

## Case 3: Top Secret-to-Secret

This section will be provided in a later release of the Framework.

### 6.7.1.6 Framework Guidance

In this section guidance is provided on the solutions that can be implemented now to perform High to Low network connections for the purposes of communication, network access and releasability.

## Case 1: Secret-to-Unclassified

### Requirement Considerations

In order to place the framework guidance in a proper perspective, this section delineates the specific security requirements being addressed and discusses issues associated with providing solutions for them.

#### a) Communication

- Secret users must be able to send and receive Unclassified electronic mail with communication partners on the Unclassified network.  
*This requirement opens the possibility of leakage from Secret to Unclassified and also the possibility of attacks being encoded in messages received from the Unclassified network.*
- Secret users must get notice of electronic mail that was sent to users on the Unclassified network but was unable to be delivered, i.e., bounced messages.
- It must be possible to send and receive electronic mail with attachments.  
*Attachments greatly increase the risk of leakage Secret to Unclassified, and the risk of attack to the Secret network, because it is generally very difficult to determine whether an attachment contains an executable.*
- Secret users must be able to participate in live Chat sessions with users on the Unclassified network.
- Secret users must be able to use collaborative technologies such as whiteboarding and video conferencing with users on the Unclassified network.
- Internet Telephony between Secret network users and Unclassified network users must be enabled as the technology becomes available.

## b) Releasability

- Enable Secret users on the Secret network to designate and push, e.g. FTP, e-mail, HTTP Post, etc.; data to the Unclassified network that is releasable to users on the Unclassified network.
- Enable Unclassified users on the Unclassified network to access the released Unclassified data using Web technology and FTP database access techniques.
- Access to Unclassified data released from a Secret network may be restricted to specific Unclassified users, groups of users, or may be made publicly available.
- The format of Unclassified data released from a Secret network may be text, video, images, audio or executable software.

## c) Network Access

- Secret users on the Secret network must be able to access the vast information resources on the Unclassified network using HTTP, FTP, Gopher, Wide-Area Information Service (WAIS), SQL, or Web Push.
- When using Web Push as a result of a previous Secret user request to the Unclassified network, Unclassified information is pushed into the Secret network from the Unclassified network

*The implications of these requirements are the dangers in retrieving data from servers. Data could harbor malicious executables. Also, information normally transmitted using the HTTP protocol might give the Unclassified servers a passive intelligence gathering capability.*

*Additional, derived requirements, which are driven by the selection of a high level architecture from Secret-to-Unclassified network access are as follows:*

- Secret users must be able to use search engines that reside on the Unclassified network. This effectively means keywords must be sent from the Secret user to the Unclassified search engine.

*The main implication of this is that data must be transmitted from Secret to Unclassified via the HTTP POST method. This method allows arbitrary data to be posted to an HTTP server. Measures must be taken to assure that Secret data is not being posted to an Unclassified server.*

- The Secret client needs to receive data of arbitrary type and format.  
*This requirement increases the possibility of attack on the Secret client. The arbitrary format of the data makes it virtually impossible to detect any undesired executable.*
- Error conditions sent by Unclassified servers must be received by Secret clients.
- The WWW interface must generate error and warning messages when it is unable to fulfill the request of a Secret client, and the Secret client must receive these messages.

## Recommended Security Policies

The security policy for the Secret-to-Unclassified connection must include statements requiring countermeasures for attacks described previously.

For passive attacks the security policy must address:

- *Traffic Analysis* - The guard shall include measures to make all network access requests coming from the Secret network anonymous.
- *Monitoring Plaintext* - Encryption shall be used for all electronic mail passed out of the Secret network. Encryption shall be used between the high workstations and all external hosts receiving data for releasability. Encryption shall be used with all Unclassified hosts that support it (for example, via SSL, IPSec). The minimum size of the encryption key shall be 80 bits.

For network-based attacks the security policy must address the following attacks:

- *Modification or Insertion of Data in Transit* - All data in transit shall have either a digital signature or keyed hash algorithms applied. These cryptographic algorithms must be deployed in conjunction with timestamps or sequence numbers to prevent replay of valid data.
- *Insertion of Hostile Executables* - Scanning for viruses and blocking applets and other executables must be performed for all data being transmitted into the Secret network.
- *Defeating Authentication Mechanisms* - Strong cryptographic authentication must be used across the enclave boundary. No Unclassified users shall access the Secret network unless it is done in accordance with the framework guidance for remote access.
- *Session Hijacking* - Continuous authentication along with timestamps or sequence numbers shall be used to prevent session hijacking.
- *Establishment of Unauthorized Network Connections* - Policy shall prohibit connections between the Secret and the Unclassified network other than those providing adequate security countermeasures.
- *Masquerading* - E-mail sender authentication and authorization to release data or to access the Unclassified network shall be handled using digital signature.
- *Manipulation of Data on the Secret Network* - This shall be handled through blocking of executables, and authentication of any users on the Unclassified network that access the Secret network remotely.

The security policy to prevent insider attacks involves procedural, physical and personnel security. The primary technical countermeasure is to implement audit and intrusion detection systems on the Secret network.

For development, production, and distribution attacks, the vendors of all commercial security products shall use approved configuration control techniques and approved distribution methods.

## Recommended Topology

The IATF recommends the topology shown in Figure 6.7-2 for the near term Secret-to-Unclassified solution.

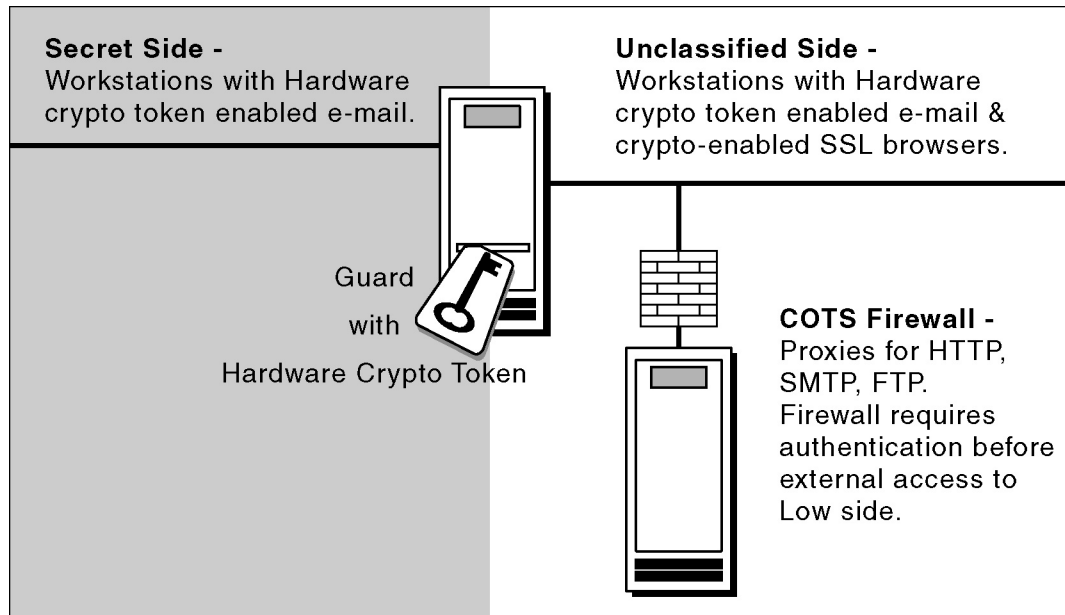
The figure shows that the only service offered between Secret and Unclassified is e-mail at this time. The guard enforces the policy for release of messages from the Secret user side. This policy can include content filtering, crypto-invocation check, release authority check, message format check, valid receiver check, message non-repudiation signature, sequence signature, and allow/disallow attachments. The policy for admittance of messages to the Secret network can include all of these elements except crypto-invocation check. The guard will be able to decrypt

copies of encrypted messages being released. However, if messages being admitted to the Secret network are encrypted, the guard will not be able to decrypt them. Consequently, the guard will not be able to filter incoming messages that are encrypted.

With minimal work, current mail guards can be modified to allow for releasability for Secret-to-Unclassified networks. It will take considerably more work to enable network access between Secret and Unclassified networks with adequate risk mitigation, because the risks of network access are quite high. The Technology Gaps section outlines a migration path to allow near term Secret-to-Unclassified capability for releasability and mid term capability for network access.

For the near term it is obvious that the guard will remain the linchpin of Secret-to-Unclassified connectivity. Many risks exist that guards will never be able to mitigate. The long-term migration should be to minimize the number of Secret-to-Unclassified connections while working to migrate toward MLS on the desktop workstation and within the servers.

The optimal solution to minimize risk is to move away from Secret-to-Unclassified and move toward MLS. MLS could be implemented on the desktop using CMWs or the Starlight Interactive Link technologies. There are several medium assurance (B2-B3) platforms on the market that are now being used as guard platforms. These could be converted to use as server platforms. Data could be separated on the network cryptographically. The technology exists for MLS; the business case has been the problem. The MLS systems that have been developed by industry have met with a lukewarm reception by government customers. Only if the government is serious about using MLS will MLS become available.



**Figure 6.7-2 Recommended Topology**

## Technology Gaps

This section addresses the near term technology advances that should be addressed to allow Secret-to-Unclassified releasability, then the mid term advances for Secret-to-Unclassified network access.

- a) *Technology Gaps for Communication* – The technology to allow communication Secret-to-Unclassified via electronic mail is readily available. However, the technology to allow Chat, Whiteboarding, Internet Telephony, and video conferencing across the network boundary is not yet available.
- b) *Technology Gaps for Releasability* – All of the capabilities needed to support Releasability are currently technology gaps. However, it is felt that Secret-to-Unclassified releasability can be accomplished within 2 years using the present solution topology shown in Figure 6.7-2. The goal is to allow users on the Secret side to submit files to the guard for downgrading. Then those files should be stored on a releasability server on the Unclassified side, making them available to Unclassified side users. They could also be made available to users outside the firewall, with the firewall and the releasability server performing authentication and controlling dissemination.

This should be accomplished by developing a releasability policy for the guard and then applying the policy to files being mailed to the releasability server. The releasability policy would likely be different than the message release policy applied to regular e-mail. The guard would recognize e-mail destined for the releasability server and would apply the releasability policy. The releasability policy will be more restrictive than the message release policy in the following ways.

- Only a very small set of users on the Secret side shall be allowed to release files to the releasability server.
- The guard shall maintain a list of this set of users and check the list upon each submission of a file to be released.
- All files submitted for release require signatures by two of the authorized individuals; one is a non-repudiation signature; the other is a sequence signature.
- Only files with specific formats of plain text or HTML shall be releasable.
- Strict audit logs shall be kept on the guard of all files sent to the releasability server.
- Released files shall be scanned for content.

The releasability server should be a COTS product that receives the files and stores them for future publication. Publication occurs when an authorized user on the releasability server unwraps the files from their signed MSP wrappers, and places them in a directory that is accessible to other users. The authorized user of the releasability server must set the appropriate permission on the published files to allow the intended users to access them.

- c) *Technology Gaps for Network Access* – There is considerably more work to be done for network access. A completely new set of filters and proxies must be developed for the guard to recognize HTTP, FTP, Gopher, WAIS, SQL, and Web Push protocols and to apply appropriate policies to these. Work is needed to develop these policies and vet



them to gain confidence that they adequately mitigate risk for network access. Elements of such a policy must include but not be limited to the following.

- HTTP Post is not allowed Secret-to-Unclassified.
- Certain fields within the HTTP protocol that identify the user making the request and the version of the browser being used must be set to arbitrary values, effectively making the Secret user anonymous.
- Executables must be blocked from entering the Secret network as Java applets or Active X controls.
- The guard shall maintain a list of Universal Resource Locators (URLs) to which access is authorized, and enforce the policy that these URLs are the only ones accessible. The guard shall perform stateful filtering of HTTP.
- The guard shall prohibit Secret users from using the FTP PUT command.
- The guard shall maintain a list of users on the Secret network that are allowed to perform network access and network access attempts using SSL.

## Case 2: Secret US-to-Secret Allied

This section will be provided in a later release of the Framework.

## Case 3: Top Secret-to-Secret

This section will be provided in a later release of the Framework.

## 6.7.2 MLS Workstation

This section will be provided in a later release of the Framework.

## 6.7.3 MLS Servers

This section will be provided in a later release of the Framework.

## 6.7.4 MLS Network Components

This section will be provided in a later release of the Framework.

# References

---

1. Reference: SSL 3.0 Specification, Netscape Communications.  
<http://home.netscape.com/eng/ssl3/index.html>.
2. Myong H. Kang, Ira S. Moskowitz, Daniel C. Lee. A Network Pump. Proceedings of the 1995 IEEE Symposium on Security and Privacy, pp 144-154. Oakland, CA.
3. Myong H. Kang, Ira S. Moskowitz, Bruce E. Montrose, James J. Parsonese. A Case Study of Two NRL Pump Prototypes. Proceedings of the 1996 ACSAC Conference. San Diego, CA.
4. Myong H. Kang, Judith N. Froscher, Ira S. Moskowitz. An Architecture for Multilevel Secure Interoperability. Proceedings of the 1997 ACSAC Conference. San Diego, CA.

**UNCLASSIFIED**

Multi-level Security (MLS)  
IATF Release 3.0—September 2000

This page intentionally left blank.

**UNCLASSIFIED**

## Chapter 7

# Defend the Computing Environment

Defense of the computing environment focuses on the use of Information Assurance (IA) technologies to ensure the availability, integrity, and privacy of user information as it enters, leaves, or resides on clients and servers. Clients are the end user workstations, both desktop and laptop including peripheral devices while servers include application, network, web, file, and communication servers. Applications running on clients and servers may include secure mail and web browsing, file transfer, database, virus, auditing, and host-based Intrusion Detection Systems (IDS) applications. Defending the computing hardware and software from attack may be the first line of defense against the malicious insider or it may be the last line of defense against the outsider who penetrates the enclave boundary defenses. In either case, defending the computing environment provides another layer of defense needed to establish an adequate IA posture.

As illustrated in Figure 7-1, the computing environment may reside within a physically protected enclave or it may be the host platform of a traveling user.

The environment includes the host or server applications, Operating System (OS), and client/server hardware. To date, the defense-in-depth technology strategy has identified the need for secure applications and secure operating systems to be implemented on

clients and servers. These security technologies are addressed in Section 7.1, Security for System Application, of this chapter. The secure applications considered are secure messaging, secure web browsing, file protection, and mission specific applications. Virus and intrusion detection software installed on host platforms is covered in Chapter 6, Defend the Enclave Boundary/External Connections.

**Security-Enabled Applications**—An application is any software written to run on a host, and may include portions of the operating system. Although there are multiple strategies for security enabled applications, this Framework emphasizes the use of open standards and Commercial-Off-The-Shelf (COTS) solutions. Continuing development of Application

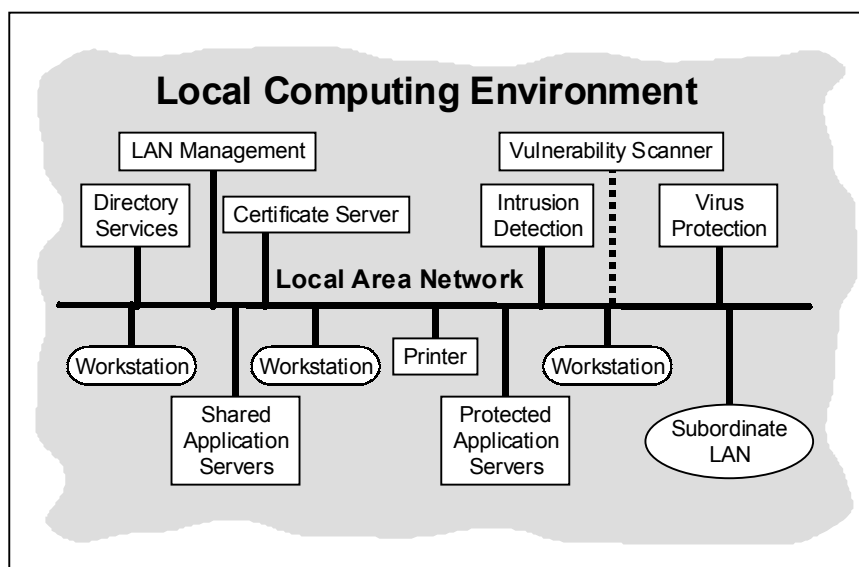


Figure 7-1 Local Computing Environment

Programming Interfaces (API) will simplify and improve the interoperability of the solutions and produce a set of standards for use throughout the Government and commercial community.

**Securable Operating Systems**—In general, the IA strategy is to provide a centrally managed, securable and securely configured operating system foundation. The vast majority of a system's life occurs after it is initially configured. Tools are needed to ensure that the initial configuration is secure, that only needed services are enabled, that vendor updates and patches are maintained, that subsequent changes retain security, and that systems are checked regularly to ensure the configuration remains secure.

**Host-Based Monitoring**—Host-based monitoring technologies include malicious software, e.g., virus, detection and eradication; detection of software changes; checking of configuration changes; and audit, audit reduction, and audit report generation. Monitoring mechanisms include tools run by users, e.g., anti-virus software, and tools managed by system administrators.

For example, administrators use network and host-based vulnerability analysis tools to verify that vendor patches are installed, detect weak user passwords, and monitor for excessive use of user access privileges. Virus protection software should be used within local computing environments.

## 7.1 Security for System Applications

This section addresses the security services that an application, in cooperation with a well-designed and maintained Operating System (OS), should provide an information system. It suggests features of both OSs and applications that provide security services, but does not reiterate well-known desirable properties or features of OSs such as the principles described in the Trusted Computer Security Evaluation Criteria (the Orange Book) or its addenda (the Rainbow series).

This section is concerned with building information systems that contain information at many levels. Everyone connected to the system is not granted access to all the information on the system. The OS provides basic separation of data. The OS protections alone, however, are insufficient for the level of separation desired to adequately secure the information. The shortfall comes in two areas: first the OSs are not written to meet any specific security policy; and second, the OSs allow privileged software to bypass or work around the operating system restrictions. Applications attempt to make up for the OS shortfalls by providing some security features and services, specifically the use of cryptography to encrypt data.

This section examines the security features and services that applications can or should provide, particularly with respect to the use of cryptography and good design practices. Several technology areas are considered:

- Network-to-network communication.
- Cryptographic Security Services and Cryptographic Application Programming Interfaces (CAPIs) that provide generic encryption, key exchange, signature, and hash functions, or higher-level security services for applications developers.
- Executable content or software download, which will consider software upgrade issues, e.g., firmware updates.
- Applications themselves that can be basic, relatively straightforward, taking advantage of security services for their functionality or extremely complex adapting basic functionality to meet a particular mission need.

In each of these technology areas the section describes specific security considerations, and specific security and interoperability concerns. These will include alternative technologies, protocols, and standards for interoperability that may be useful to those building complete and real systems.

This section generally follows the format established in other sections: Target Environment, Consolidated Requirements, Potential Attacks, Potential Countermeasures, Technology Assessment, Cases, and Guidance. The concerns for electronic mail (e-mail), distributed databases, file encryption, Internet phone, and Web-based applications have similarities, but also differences because of use, technology, standards, and more. In the major sections, the common aspects of application-level security are considered. In the technology assessment section, additional, more application specific information is supplied.

## 7.1.1 Target Environment

The environment for user- or application-layer security is generally considered to be a workstation (laptop, desktop, etc.) connected at least part of the time via a network to sensitive information servers. Additionally, the information on the servers (and on the workstation) may need protection even from other personnel or workstations privileged with access to the network resources. Further, the section assumes that the environment is the “application space” where user’s and applications operate on information that has value. Physically, this environment applies anywhere within the Global Information Infrastructure (GII) that a particular application might send, store, retrieve, or destroy sensitive information. It typically embodies the elements of a three-tier model: the client, the business process, and the databases that serve a particular process.

### 7.1.1.1 Environment Aspects Outside Direct Application Control

There are several aspects of the environment that are outside the direct control of the application that must be true prior to installing an application in order for the application to provide security to data being processed on the workstation. These aspects are fundamental to the environment in which user- or application-level security services would be applied. These apply equally well at both the client and server, or on all platforms within a distributed application environment. These aspects include the following.

- Workstation hardware operates correctly and its integrity is intact from its trustworthy developer, e.g., the hardware is free from malicious code.
- The OS operates correctly and its integrity is also intact from a trustworthy developer.
- Cryptography, when invoked, works correctly and as intended with the underlying cryptographic algorithms being secure and properly implemented.
- Cryptographic key material is provided by a trusted key infrastructure via trusted or secure means and correctly stored and protected from unauthorized access.

Applications provide security for information or data for one of two primary purposes: to limit the movement of, or access to, information within the system and to prepare information so it can cross system boundaries. In limiting access, the application generally encrypts data, although alternate methods include users logging into an application with some associated access rights determined by an administrator, or limiting access to storage media, e.g. removable hard drives. In preparing information to cross system boundaries, encryption may be used, but labeling, signing, scanning (for “dirty words”) or other methods may also be used to prevent unauthorized data from leaving the system.

### 7.1.1.2 Boundaries

Although boundaries are normally thought of at the edges of enclaves, e.g., firewalls fronting an enclave or a guard between security levels, applications form a type of boundary around themselves through mechanisms designed to prevent unauthorized access. OSs and applications can take steps to prevent, detect, and audit unauthorized data modification attempts.

Even if firewall or guard boundaries are providing security or assurance to the overall system, the application's obligation to protect sensitive information is not diminished.

### 7.1.1.3 Applications Environment

The environment for applications is considered to be a well-managed UNIX or Windows NT OS, managed by knowledgeable system administrators, using security principles and practices in a documented networked environment, using all known system patches for security, and following good management practices to maintain a system information policy. Most applications will be commercial in nature, i.e. the foundation will be commercial packages, but increasingly, the application will require customization to fulfill a specific business process need. The customization may take many forms, and the coding language used by the custom applications will have an impact on the security of the resulting system. Highlights of these coding languages follow.

C and C++ are widely regarded as portable languages that allow applications to move across platforms. Compilation options, and non-standard terms may create debugging problems.

Common Gateway Interface (CGI), Practical Extraction and Report Language (PERL), JavaScript, Microsoft Macro Language, and similar scripting languages are very powerful, with cross platform capabilities. Their power makes these languages good targets for hacking attacks as they support both local and network capabilities.

JAVA is billed as cross platform, but like C and C++ great care must be used in writing actual code to ensure cross platform capabilities. The JAVA language is somewhat unique in having a security model (the sandbox) but the concept greatly limits the usefulness of some applications. Efforts to expand the sandbox are making JAVA more like ActiveX, providing more capability, at greater risk and with some user trust of the software provided through interfaces and signed code.

ActiveX is a Microsoft unique language/capability for distributed custom applications. ActiveX is very powerful, but the security model is a fairly simple model based on signed code with authenticated signatures. The flexibility is a concern to many security professionals.

There are other languages with other concerns available on various platforms. The four cases of software application being considered in this section are generally assumed to be well-written code from developers lacking evil intent. The environment assumes that the vendor code functions as intended, without bugs.

### 7.1.1.4 Operating Systems Environment

This overall section of the Framework is focused on the security services that applications could provide to protect data that the applications manage and manipulate on behalf of the workstation users. This data may be intended for private, narrowly shared, or widely shared consumption. Typically, an OS provides users with the ability to share the various hardware resources of a workstation or computer system. The OS virtualizes and manages access to memory, disk drives, data ports, and other hardware resources. This management separates users so that one user's memory space cannot be read by another user's process. The OS management also allows for some degree of portability so that software code written on one



machine may be ported to another machine with less difficulty than if all code directly called the hardware.

The Orange Book and many other publications detail the security functionality an OS should provide. This has been done from the standpoint of trusted software, where no cryptographic functionality was present. Only recently have people begun to analyze the value and functionality that cryptography can add to a trusted OS and therefore might provide to applications. This second phrase is important because a trusted OS provides protection and integrity to the cryptographic functions that an application cannot truly provide on its own.

An OS provides several basic mechanisms and capabilities to support information system and application security. The requirements for these capabilities have been widely written about, in the Orange Book, in the Common Operating Environment (COE) requirements, and in the Common Criteria. A specific set of requirements for OSs is being captured in Common Criteria Protection Profile Format through the Defense-wide Information Assurance Program (DIAP) to document requirements for Protection of Host Computer OSs (Clients and Servers).

### 7.1.1.5 Standards and Protocols for Providing Security to System Applications

Efforts at standardizing security features and services have attempted, as a primary goal, to specify algorithms, formats, protocols, configurations, etc. If standardization is accomplished, the common security services (Section 4.4, Important Security Technologies) can be provided to protect against the universe of threats (Chapter 3, System Security Methodology) with the maximum level of interoperability (Section 4.6, Interoperability Framework).

From an environment standpoint, this Framework emphasizes the importance of utilizing open standards and Commercial-Off-The-Shelf (COTS) solutions. Commercial implementers are becoming more and more dedicated to generating and implementing open standards. Such standards enable multiple independent implementations to be built and interoperate. Public disclosure of the details of security protocols and algorithms is being demanded by the security community at large so that these standards may be evaluated and tested to an appropriate level of assurance.

The term “standard” is used quite loosely in this Framework. The term is meant to include any standard, technology or product initiative that could evolve into a standard. Standard can encompass national, international, Department of Defense (DoD), federal, allied, and commercial standards. The term may also include various initiatives that are not specifically standards activities but that may become so or that may evolve into de facto standards. This Framework primarily addresses standards relating specifically to security, but may also include other standards that affect interoperability or system infrastructure. In many areas, security is simply an element of a broader standards activity.

Specific examples of standards and protocols of interest include (see also Section 4.4, Important Security Technologies):

- Application Layer.
  - Secure HyperText Transfer Protocol (S-HTTP).

- Object Management Group's Common Object Request Broker Architecture (CORBA).
- W3C XML Transfer Protocol.
- Secure File Transfer Protocol (S-FTP).
- Secure Electronic Transactions (SET).
- Message Security Protocol (MSP).
- Secure/Multipurpose Internet Mail Extensions (S/MIME).
- Transport and Network Layer.
  - Transport Layer Security (TLS).
  - Secure Socket Layer (SSL ver 3.0).
  - Secure Shell (SSH).
  - Internet Protocol Layer Security (IPSec).
- Data Link Layer.
  - Point-to-Point Protocol (PPP).
  - Serial Line Internet Protocol (SLIP).
- Security Management Infrastructure.
  - Internet Engineering Task Force (IETF) Public Key Infrastructure.
  - IETF Simple Public Key Infrastructure (SPKI).
  - IETF Domain Name System Security (DNSSEC).
- Data Labeling.
  - National Institute of Standards and Technology (NIST) Federal Information Processing Standard (FIPS) 188 Standard Security Label.
  - Institute of Electrical and Electronics Engineers (IEEE) 802.10g Secure Data Exchange (SDE) Security Label.
  - IETF Internet Security Label.
  - International Organization of Standardization (ISO) SC-32 Security Label.
  - Military Standard (MIL STD) 2045-48501 (Common Security Label).
  - SDN.801 Reference Security Label.
  - ISO MHS X.411 Security Label.

## 7.1.2 Consolidated Requirements

At the highest level, security requirements for applications can be divided into two areas: functionality and assurance.

Functionality is the simpler area. The application security functionality requirements are simply a list of what security functions the application must supply in order for the information on the system to be protected. Functionality requirements can usually be specified and tested in an objective fashion. Requirements for application layer software span a broad range of functionality. This range covers local application activity and extends to the many different approaches to communication and collaboration between users. The difficult question for application security functionality requirements is where should the requirements be levied, in the OS or the ap-

plication program. Common, widely used functionality, e.g., file system access control, belongs in the OS. Specialized functionality belongs in applications.

Assurance is a more subjective requirement. Assurance is a measure of confidence that the security features and architecture of an information system accurately mediate and enforce the security policy. Assurance requirements provide confidence that an application meets its security goals. Desired assurance includes that the functionality requirements were properly designed, implemented, and tested. There are many different approaches to assurance. Process assurance requires the software developer to adhere to a specified software-engineering life cycle. Product evaluation investigates and assesses the design and development of a product before approving it for use. Black-box testing examines outputs for specific inputs. Mathematical analysis can be used to formally model the security policy and assesses cryptographic strengths. Assurance provides increased confidence about the “goodness” of a security product’s security features. It does not make the product more secure. A product with little or no assurance may still be very secure. One just doesn’t have a significant confidence level and accreditors may refuse to accredit the application for the desired use.

### 7.1.2.1 Functional

At the highest level of abstraction, the comprehensive functional requirements are very general in nature. As more specific applications are considered these functional requirements are refined to more specific goals. High level functional requirements include the following.

- The application shall be user-friendly with well-documented user interfaces.
- The application shall use a correct and efficient implementation of backend processing.
- The application shall support standards and implementation with standards-based Application Program Interfaces (APIs).
- The application shall provide privacy and integrity protection of user and system data.
- The application shall authenticate the user to provide accountability.
- Management of configuration information should be centralized where possible and supported by secure remote management when necessary.

### 7.1.2.2 Interoperability

Interoperability requirements at the application layer are supported by interoperability of lower level network activity. There are also interoperability requirements on application level data formats that can be supported by compliance to appropriate standards. In instances where two systems are using different security technologies, interoperation of these security technologies is required for the systems to achieve data interoperability.

### 7.1.2.3 Anticipated Future Requirements

It is desirable that the security solutions are capable of evolving to higher data rates, to different paradigms of use, and that they be adaptable to alternative means of access.

## 7.1.3 Potential Attacks

The focus within this category is on attacks directed against system applications. Five classes of attacks introduced in Section 1.3.5, Nature of Cyber Threats, are concerns for the security-enabled applications. There are also additional specific attacks that, although once identified will fall into one of the four attack categories, can only be identified and countered at a lower design level. These are categorized as lower-level attacks. Details of these additional threats to application security as well as the four classes of attacks previously discussed are provided here.

### 7.1.3.1 Active Attacks

Protocol exchanges between clients and servers are common threads in application security. These protocols may have security as their immediate concern (authentication protocols) or they may provide application functionality with the assumption that security is already in place. Many forms of spoofing and network connection hijacking have been observed in the field. Additional analysis has also identified vulnerabilities in security protocols that were widely believed to be correct.

### 7.1.3.2 Passive Attacks

The nature of passive attacks can vary greatly. Information collected may be clear-text or encrypted. Encrypted information may be subsequently subject to various forms of cryptanalysis. Passively captured information may be used to support network replay attacks.

### 7.1.3.3 Insider Attacks

Attacks launched by trusted users inside an enclave are considered “insider attacks.” Insiders can be employees, contractors, service providers, or anyone with legitimate access to a system. The term “cleared” insider refers to a person who holds a clearance and has physical or administrative access to classified AIS.

Protecting against and detecting malicious behavior by insiders is one of the most difficult information assurance challenges. Both technical and procedural countermeasures can reduce the risk, but to be effective, technology and procedures must complement one another. Countermeasures to this form of attack lie partly in the realm of background checks and physical security. Limits can be placed on each individual’s authorized privileges. The application and the security features it provides can also partly counter these threats with features such as audit, two-person administrative requirements, and covert access prevention and detection.

### 7.1.3.4 Distribution Attacks

The risk of malicious code in commercial application software is difficult to quantify and similarly it is difficult to judge the value of applied countermeasures. For mass-produced office application software, which can be obtained from many sources and compared, the risk of hidden attacks in the software seems small. As one moves to custom applications created for specific

security-conscious organizations, the malicious software risk must be considered more carefully. The defensive options include review and control of the source code and/or security requirements on the application producer's software development process.

### 7.1.3.5 Lower-level Attack Analysis

Poor protocol specifications may enable one source of lower-level attacks. Careful analysis of the specifications of protocols such as Transmission Control Protocol (TCP) and Server Message Block (SMB) can identify opportunities for attacks that compromise information or deny service. Included in the draft SMB protocol specification is its security limitations.

Beyond the protocol specification, the details of specific implementations can enable attacks. For example, some implementations that use a simple predictable algorithm to generate initial sequence numbers are susceptible to a well-known spoofing attack.

Another common group of lower-level attacks exploit the failure of application code to do memory bound checks or other error analysis on data provided by external sources. Buffer overflows and other tricks can then be used to cause malicious remote command execution.

## 7.1.4 Potential Countermeasures

Information systems can be susceptible to attacks at many levels. Countermeasures must span a similar range. There are countermeasures that apply to the entire system. There are also application specific countermeasures. At the lowest level countermeasures must respond to implementation specific attacks. Countermeasures must continually be improved to counter more sophisticated attacks. The ultimate goal is for the countermeasure to become so sophisticated that the cost of mounting the attack exceeds the value gained from a successful attack. The threat to the information system is reduced when the rational attacker discovers the reward based upon the effort of the attack.

Countermeasures are enabled through the use of various security services. The security services may use cryptographic mechanisms to provide their functionality. These cryptographic mechanisms include Public Key Certificates, Key Exchange (Public Key Cryptography), Data Encryption (Private Key Cryptography), Digital Signatures, and Cryptographic Hashing. Chapter 8 (Supporting Infrastructures) of this Framework is devoted entirely to using cryptographic services, keys, certificates, and the key infrastructure to support these security services. The chapter concerns itself with Public Key Infrastructures/Certificate Management Infrastructures/ Security Management Infrastructures (PKI/CMI/SMI), and the capabilities, security considerations and policy that pertain to these. Functionally, the PKI/CMI/SMI is intended to provide authentication that a certificate is tied to a unique entity, secure distribution of certificates and private key material, wide distribution of public key material, and notification of compromised and revoked certificates or key material. A variety of technical and policy measures counter various attacks and security concerns related to key management. These are detailed in Chapter 8 and various public forums.

Details of the security services and the countermeasures they provide follow.

### 7.1.4.1 Access Control (Authorization)

Access control is the process of granting access to information system resources only to authorized users, programs, processes, or other systems. Controlling access can be based on identification and allotted roles, roles alone, user name, group membership, or other information known to the system on a secure basis.

A well-managed Windows NT or UNIX OS can provide a degree of basic access control. This access control limits user access to specific resources and privileges.

Controlling who can use an application protects the data the application is designed to maintain on an exclusive basis. Those that intend to alter the information or add some type of malicious process are foiled because they are denied any capability to use the application or access the data via a separate path. What is of particular importance is controlling who is allowed to enable or disable (i.e., turn on or turn off) the security features that may be built into the application or to change the privileges of users or programs.

Secure applications that process data must be aware of their role in managing access to that data. That includes knowing who is attempting access, mediating access according to some processing rules, auditing user actions, and managing where (access to printers in particular locations) or how (encrypted channels such as SSL) data is sent. Access control may be managed solely by the application, or it may use OS functionality for assistance. An example of the latter might be a database that uses OS controls on files (user/group/world read/write privilege) by putting different classes of information into different files with different access privileges in which none of the data can be directly accessed by the users.

### 7.1.4.2 Identification and Authentication (I&A)

I&A is the process of identifying the user who is trying to access a system, and authenticating the identity of that user. It is this combination that provides accountability. When used with effective access control, the more uniquely the user can be identified and the more assuredly this identity can be authenticated the more secure the system.

The user's identity can be assured by requiring the user to identify himself with something he has (e.g., an identification badge or hardware token). His identity can be authenticated by requiring the user to provide something he alone knows, e.g., a password or Personal Identification Number (PIN) and something he is, e.g., a fingerprint, retinal scan, or other biometric.

Electronic or digital signature can also assure user authentication. Identification can be based on a public key certificate, an electronic certificate signed by an issuer. This provides a unique digital identity for the holder of the certificate. Validation of the certificate chain is part of the authentication process. The certificate issuer provides the authentication of the identity based on possession of the issued certificate.

### 7.1.4.3 Data Integrity

Data integrity is the state that exists when data is maintained as intended and has not been exposed to accidental or malicious modification. Data integrity is a separate function from data

encryption, although some encryption algorithms can be used to prove that integrity has been maintained.

An OS and an application can work together to protect data from modifications. The OS can provide integrity on its files. Files can be saved, opened, modified, and closed by applications, with the assurance, from the OS, that the information on the files has only changed if an authorized application changed the information.

The application and the OS can provide additional integrity through the use of a cryptographic hash function. Each entry is mathematically hashed producing a unique value for any given entry. Verification of the hash guarantees the data integrity. A digital signature applied to the hash value allows for the authentication of the hash value and who applied it. It is important to note that hashing is a one-way function. That is, the hashing algorithm cannot be reversed to reconstruct the data from the hash value.

### 7.1.4.4 Data Confidentiality

Data confidentiality is a property asserting that information is not disclosed to unauthorized entities or processes. Access control mechanisms support data confidentiality in information systems by controlling access to the system's resources.

From an application standpoint, confidentiality is especially important when the application is not running. Without the OS or application providing access control, the data in storage is especially vulnerable. Data in transit, outside the direct influence of its generating application, is also vulnerable. Encryption is useful in either case. Both applications and OSs can provide encryption of stored data and data in transit. The measure of data confidentiality is directly related to the algorithm used to encrypt data and the protection of the key used for encryption.

### 7.1.4.5 Availability

Availability of applications and data is critical to the security of information systems. Data that is inaccessible might as well not be there. Likewise, applications that fail to work are useless. The operating system and applications should be designed to withstand failure in either the OS and in the applications. Most UNIX systems and Windows NT have error handling routines and fault isolation, providing better availability of the OS in the event of application failures. Applications should be designed and tested to ensure that they do not fail, particularly under extreme conditions. The robustness of an application cannot prevent problems when the underlying OS fails, or when external network components (guards, firewalls, routers, cable) fail.

### 7.1.4.6 Non-Repudiation

Non-repudiation is a property by which the recipient is assured of the originator's identity and the originator is provided with proof of delivery, so that neither can later deny having processed the data. Non-repudiation counters man-in-the-middle and spoofing attacks. One way to achieve non-repudiation is with digital signatures and auditing. User authentication assures the recipient of the originator's identity. Before transmitting, the originator signs the data with an algorithm that incorporates parameters unique to the originator. Verification of this signature verifies the originator's identity. Auditing makes a complete record that can serve as evi-

dence and protects the record's integrity. For proof of delivery when sending data, the originator requests a signed receipt. The recipient signs the receipt with an algorithm that incorporates parameters unique to the recipient. Verification of this signature verifies the recipient received the data.

Since non-repudiation depends on an identity often contained in a public key certificate that can become invalid, it is important that a "trusted" third party be able to establish the validity of the certificate. It must be possible to prove the validity of that certificate at the time of the original communication and that the authentication be recorded in the audit trail.

### 7.1.4.7 Auditing

Both the application and the OS can audit certain actions taken by users and software acting on the OS. An application might track when a user enters data into a database, and information related to the data or its position in the database. An OS might track which users initiate a process, or attempt to access particular files. Auditing is primarily an after-the-fact activity that supports information forensics activities, and intrusion detection. Intrusion detection tools have been developed to detect intrusions into a computer or network through observation of the security logs or audit data. These tools can be an integral part of the OS or the application. Intrusion detection tools can also be separate software added to a system for the express purpose of detecting intrusions. See Chapter 6 ,Defend the Enclave Boundary/External Connections, and Section 7.2 for an in depth discussion of intrusion detection.

Auditing is a protective measure only in the sense that knowledge of auditing may deter some activities that threaten information systems. Auditing is much more useful in detecting questionable activity, and reacting to such activities.

When developing an application, the developer should make explicit use of OS audit capabilities, and plan for system administrators' or other security professionals' use of the audit data. One of the overarching technology gaps today is the availability of useful audit tools.

## 7.1.5 Technology Assessment

The three technology areas cryptographic security services, applications, and software upgrade will be considered separately.

### 7.1.5.1 Cryptographic Security Services

Two pieces are needed to provide cryptographic security services to applications. First, and foremost, some type of cryptographic algorithm must be available for the application to use. It is not the intent of this Framework, however, to assess specific algorithms. This Framework will assess the medium, token, on which the algorithm is presented to the application for use. The algorithm on the token is presented to the application, through a CAPI.



### 7.1.5.1.1 Cryptographic Tokens

Stand-alone cryptographic devices met the security needs of the past. Confidentiality was the security service of choice, which was implemented with link encryption, one device servicing many users. The need for security services beyond confidentiality has arisen with the growth of network technology. One such needed service is I&A—the need to specifically name users and have assurance that the persons associated with those names are who they claim to be. Cryptographic technology has progressed as well, in both size and cost, and can now provide security services on a personal basis. Any user can have their own personal cryptographic device, or security token, which is uniquely their own. Tokens can also provide data integrity and non-repudiation services through hashing and digital signature algorithms.

Using a personal security token that implements public key cryptography enables each user to have a private key that is unique to them. This key can then be used as the basis for the security services of non-repudiation, and I&A. One way to accomplish this is to use the keys to create digital signatures on messages. Upon receipt of such a signed message, the recipient can verify the sender's digital signature and therefore accept that the message is truly from the user who claimed to send it. Tokens can come in different forms—from Personal Computer Memory Card International Association (PCMCIA) cards to smart cards to even software. Each implementation offers advantages and disadvantages.

#### PCMCIA Tokens

A PCMCIA Security Token can offer a full suite of security services in a portable format. Board real estate allows room for sizable memory parts such as Random Access Memory (RAM) and Electrically Erasable Programmable Read Only Memory (EEPROM) or Flash EPROM providing ample memory for complex or multi-function firmware and certificate storage. Since it is a hardware token, it can also protect secret values reasonably well and still leave room for additional physical tamper protection mechanisms. On the down side, PCMCIA cards require PCMCIA card readers which, although they are prevalent in laptop computers, are not found too often in desktop computers. The added expense of purchasing a card reader for every desktop workstation is definitely a disadvantage of the PCMCIA token.

#### Smart Card Tokens

Smart cards offer the same portability as the PCMCIA token but at a reduced cost. They still require special readers but a smart card reader is much less complex than a PCMCIA reader and thus less expensive. Some manufacturers are incorporating smart card readers into their computer keyboards. One significant concern in the smart card arena is that of data throughput. The defined interface is just too slow to support confidentiality services for any but the least demanding applications. Confidentiality would normally be relegated to software running on the workstation, which can reduce the assurance of this service. I&A, non-repudiation, and data integrity would still remain implemented in the hardware on the smart card.

#### Software Tokens

Software tokens offer the cheapest solution but also the least assured solution. Implementation in software allows for quick distribution, ease of updating, and a solution that can meet the

needs of most users without the need for a special reader. When the security solution calls for minimal features and assurance and when cost is a major consideration, software tokens could be the answer. There is a price to be paid with software, though. That is its assurance. Software tokens will execute on untrusted workstations running untrusted OSs that make them ultimately vulnerable to bypass, modification, or even replacement. Systems processing highly sensitive information should not rely solely on security services provided by software tokens.

### 7.1.5.1.2 Cryptographic Application Programming Interfaces (CAPI)

As application developers become aware of the need for cryptographic protection, they add “hooks” to access the cryptographic functionality developed by others. These “hooks,” at the lowest level (sometimes crossing into the OS and almost always within what would be called middle-ware), are the CAPIs. As CAPIs mature and gain sophistication, the benefit derived from these standard interfaces increase. Applications that utilize a standard CAPI can access multiple cryptographic implementations through a single interface. This helps to minimize life cycle implementation efforts. Likewise, cryptographic modules that are built to a standard CAPI can be accessed by a greater number of applications, increasing reusability.

There are numerous efforts currently under way to create CAPI standards. They range in scope from very generic security support like that found in Generic Security Services (GSS)-API to an interface more involved in the direct control of the cryptographic token like that found in Public Key Cryptographic Standards (PKCS) #11. A number of these CAPI efforts are receiving a great deal of support as applications and cryptographic modules are being written to use them. While a single CAPI standard usable by all applications would be ideal, multiple CAPIs are required to support the broadest range of applications and cryptographic modules.

CAPIs are intended to provide these features:

- Interface between cryptography and applications.
  - Facilitate the development of new security enabled applications.
  - Application does minimal cryptography processing.
- Application independence.
  - Support broad range of application types - store and forward and connectionless.
- Module independence.
  - Support entire range of hardware and software tokens.
- Algorithm independence.
  - Support broad range of current and future algorithms.
- Functional completeness.
  - Provide comprehensive security services.
  - Facilitate cryptography export policy.

**High Level – GSS-API**—The GSS-API and the extensions for Independent Data Unit Protection (IDUP) support cryptographically unaware applications. These Microsoft's Security Service Providers (SSAPIs) provide a high-level interface to authentication, integrity, confidentiality,

and non-repudiation (IDUP only) services. The application merely indicates the required security services and optionally the quality of protection (QOP) for the per-message services. GSS-API was designed to protect session-style communications, e.g., File Transfer Protocol (FTP) between entities. IDUP-GSS-API does not assume real-time communications between sender and recipient, and protects each data unit, e.g., files or messages, independently of all others. IDUP-GSS-API is therefore suitable for protecting data in store-and-forward applications. These specifications were developed within the Common Authentication Technology (CAT) group within the Internet Engineering Task Force.

**Mid Level - CDSA, MS SSAPI**—The Common Security Services Manager API (CSSM-API) is the heart of the Common Data Security Architecture (CDSA). CSSM-API offers a robust set of security services to include: cryptography, certificate management, trust policy, data storage, and optionally key recovery. CSSM-API also has the capability to provide integrity services via the Embedded Integrity Services Library (EISL) and to support auditing services. CSSM-API was developed at Intel Architecture Labs and is approved as a standard within the Security Program Group (SPG) of the Open Group (the result of the X/OPEN and the Open Software Foundation merger). While CSSM services like certificate management, trust policy, and data storage fit logically at the middle level, the actual CAPI calls (their Cryptographic Service Provider Interface (SPI)) are more low level like Cryptoki. For instance, CSSM-SPI supports user authentication and administrative control of tokens.

SSAPI is modeled after the GSS-API, though with more of a “Windows” style. It provides mutual authentication, message privacy, and message authentication. It is connection oriented, thus it is used for such protocols, defined by Microsoft as “SChannel”, SSL and WinPCT. It also supports other mechanisms: NTLM, DPA, and Kerberos.

**Low Level - Cryptoki (PKCS-11), Cryptographic API (CryptoAPI), CI Library**—PKCS #11 - Cryptoki is an OS independent abstract token interface that defines the arguments and results of various algorithms. Cryptoki also specifies certain objects and data structures which the token makes available to the application. Cryptoki interfaces directly to cryptographic tokens, and is thus the logical place for functions that allow user authentication (e.g., logon or PIN entry) and administrative control of the token. Cryptoki is appropriate for use by developers of cryptographic devices and libraries. Cryptoki was developed by RSA Labs and is a member of their family of PKCS. Continuing development of Cryptoki is accomplished by the PKCS #11 workshops sponsored by RSA Labs and held annually for all interested parties.

The Microsoft CryptoAPI supports cryptographically aware applications. As a service suite provided by the Windows NT OS, CryptoAPI provides extensive facilities for utilizing both hardware and software cryptographic modules, called Cryptographic Service Providers (CSPs). CryptoAPI was developed by Microsoft and therefore has not been subjected to any formal standards process. However, the authors did consult with various Government and corporate customers while developing the CryptoAPI specification. Applications using CryptoAPI can take advantage of default features of the interface to reduce their cryptographic awareness requirements, or they can exert full control over algorithms, keys, and modes of operation.

The FORTEZZA® Cryptographic Interface (CI) Library was initially developed as the interface between the FORTEZZA® PCMCIA card and applications wishing to use the security features associated with the National Security Agency's Multilevel Information Systems Security Initiative (MISSI) program. The CI Library is now being adapted for both Smart Card and Software token implementations of FORTEZZA®.

## 7.1.5.2 Applications Requirements

Applications are generally useful for exchanging information among multiple people within a specific system, or between information systems. The applications discussed here are characterized as either “basic” or “mission” applications. Basic applications form the basis from which other applications can be devised. These include messaging, web browsing, and file protection. For mission applications, basic application functionality has been adapted to meet a particular mission need. Examples of these include databases, collaborative computing applications, and electronic commerce systems. Because information is being transmitted, the need for standard, interoperable, and secure applications is critical. While many applications are mature, most do not support the broad range of security services.

### 7.1.5.2.1 Generic Requirements

Applications encompass a wide range of functions. In this section, the generic requirements that apply to all secure applications are provided.

**Functional**—The functional user requirements for the secure transmission and storage of information include:

- The ability to exchange information with other users while maintaining data confidentiality and integrity,
- A system support for broadcasting information from one user to many users,
- The ability to determine and verify the originator of the information,
- An indication that security features are enabled, and
- Minimal impact on the user.

**Interoperability**—Interoperability is needed between user clients, between clients and servers, and among clients, servers, and security management systems. Requirements for interoperability include:

- The use of standardized user-to-user, peer-to-peer, client-to-server, and management protocols,
- Format and encryption algorithm negotiating techniques, and
- Standard for exchanging security credentials among clients, servers, and security management systems.

**Anticipated Future Needs**—The evolution of applications should allow for supporting multiple standards and negotiating the standard to use among components. Additionally, security functions should evolve into a form that allows these functions to be easily used by applications.

### 7.1.5.2.2 Tailored View of Attacks and Countermeasures

Attacks that can be launched against an application include the following.

- *Social engineering* – The attacker fools the user into disclosing password and then loads malicious software, e.g., virus, and applet; disclosing protected data; or signing bogus data.
- *Exploit data execution* – The attacker gets user's system to download and execute malicious software.
- *Masquerade as authorized user or server* – For example, a web server pretends to be another server, providing the user with bogus information and/or fooling the user into providing sensitive information.
- *Exploit system-application or OS software* – For example, the attacker grabs the decrypted information of the target user via the OS.
- *Data modification* – Data is changed before reaching its destination.
- *Deny service* – For example, the attacker add a user's key to the key compromise list.
- *Equipment modification or theft, software modification from the console*
- *Traffic analysis* – The attacker determines how much information is sent from originator to recipient, but not the actual information. For example, a dramatic increase in information between parties could be indicative of a military operation in the early stages of execution.
- *Cryptographic analysis*
- *Message replay* – The networking software may not filter a replay of a message.
- *Distribution attacks* – the attacker copies or modifies software or keying material during distribution.
- *Valid user performing inappropriate actions.*

The countermeasures to these risks are separated into countermeasures to be provided by the system and its support services and countermeasures specifically associated with the application. Every link in a system is a potential point of vulnerability. A compromised link can allow denial of service or compromise of data (e.g., intercepted or modified). System provided countermeasures can be characterized as those that must be enforced at all points along the path that data travels. Application provided countermeasures must be enforced with specific mediation. This mediation must not be bypassed.

#### **System Provided Countermeasures**

- System integrity.
- Physical protection of the hardware.
- Basic OS security.
- Backup and recovery procedures.
- Anti-virus software.
- Intrusion detection software.
- Firewalls.

- Physical protections including the protection of media containing application software and application data while it is not stored on-line, e.g., key repository.
- Appropriate procedural security.

#### **Application Provided Countermeasures**

- Confidentiality of the user information as it passes from the originating user via the server(s) to the destination user(s).
- Confidentiality of the keying material information.
- Confidentiality of the originator and recipient.
- Protection of the application software's integrity during distribution.
- Integrity of the information, key material, source and destination information, and any other control information.
- User I&A of users and administrators, system I&A of clients and servers.
- Access control as appropriate for the particular application, e.g., access control on the mail server.
- Auditing of the user, administrator, client and server as appropriate for the particular application.
- Non-repudiation may be required, e.g., need to trust that the creator of a key pair is whom he claims.
- Key and/or information recovery.

### **7.1.5.2.3 Basic Application Technology**

This section looks at the basic technology used by each of the basic applications: secure messaging, secure web browsing, and file protection.

#### **Secure Messaging**

E-mail systems provide a means of transferring messages from a source client, via a mail server, to a destination client. Source and destination addresses are sent to the mail server and used by the server to determine the proper routing. Protocols used to support e-mail include the Simple Mail Transfer Protocol (SMTP), the Post Office Protocol (POP), X.400, and Internet Mail Access Protocol (IMAP). Minimal security services are provided.

Most of the commercial secure e-mail clients available today are add-on packages or standards that build upon the features of the existing e-mail client software. These include products built to the Message Security Protocol (MSP) standard, the draft IETF Secure Multipurpose Internet Mail Extension Version 3 protocol standard, and the draft IETF Open Pretty Good Privacy (OpenPGP) standard. Other commercial e-mail systems use an integrated approach where both the client and the server participate in providing the secure e-mail services. Examples of these products are the Microsoft Exchange Mail Server and the WorldSecure Server. These secure e-mail packages and standards provide:

- Encryption of message content and attached files on the client system for both local storage and transmission.

- Encryption scheme minimizing required communication bandwidth and disk space.
- Integrity of message content and files provided through encryption and/or digital signature.
- Integrity of e-mail client software.
- Authenticity assurance of message content and files through the creation and verification of a digital signature.
- Availability as directed by priority delivery classifications.
- Non-repudiation of act of sending or act of reception.
- User to e-mail client system authentication support, e.g., a hashed pass phrase.
- Key management support, i.e., key generation, distribution, deletion, destruction, and revocation.
- Certificate management support, i.e., certificate requests, distribution, deletion, and verification.

There is a wide range of possible architectures for the Key Management Infrastructure (KMI)/PKI needed to support secure messaging. Possibilities range from a user having complete control over key generation and distribution to a hierarchical architecture involving a complex CA. KMI/PKI is discussed in detail in Chapter 8.1 (KMI/PKI).

## Secure Web Browsing

A web browser provides user access to a web server for the purpose of viewing, retrieving, and sometimes uploading information stored on the server. Web browsers also allow for scripting, the downloading of small applications or applets (often referred to as “mobile code” or “active content”; examples are Java applets and ActiveX controls), forms, and other specialized capabilities that support a visual and interactive interface to information.

Current commercial web browsers implement a number of security mechanisms. The Secure Sockets Layer protocol is the primary method by which transmission security is provided in Web-based products today. The concept of signed objects is one method for dealing with mobile code. Down-loadable code is signed by a trusted third party as an indication of its origin and integrity. The mechanisms implemented by secure web browsers provide:

- Encryption of the HyperText Transfer Protocol (HTTP) transaction between browser and server, including any authentication information, using the SSL secure channel.
- Encryption of the private portion of the asymmetric key stored on the user’s system.
- Integrity of the HTTP transaction between browser and server using SSL’s message authentication code.
- Integrity of the user system by preventing or validating the downloading of mobile code that could potentially damage the system, e.g., disabling the use of JavaScript.
- I&A of the message origin using SSL’s message authentication code.
- I&A of the server using server-side certificates that have been signed by a trusted third party.
- I&A of the user using either passwords or client-side certificates.

- Logging of the browser domain name and IP address, the accessed web page Universal Resource Locator (URL), URL perimeters, and access time on the web server whenever a document request is received; logging of exception errors when they occur.
- Logging the browser application program used, the cookie content, and the user name and e-mail address (if provided by the user to the web browser); key management support, i.e., key generation, distribution, deletion, and destruction.
- Certificate management support, i.e., certificate requests, distribution, deletion, and verification.
- Access control based on authenticated identity.
- Code check based on byte code security check.

There is a wide range of possible architectures for the KMI/PKI needed to support secure web browsing. Possibilities range from a user having complete control over key generation and distribution to a hierarchical architecture involving a complex certificate authority (CA) structure. KMI/PKI is discussed in detail in Chapter 8 (Supporting Infrastructures).

## File Protection

File encryptors protect information in the computer in the event of unauthorized physical access. This is accomplished through encryption of the stored information. There are two basic types of file encryptors: one that the user selects specific files to encrypt, and one that automatically encrypts all information that isn't currently being processed in some manner. The former can be used to securely transfer files as attachments or to protect critical information stored on floppy disk, Compact Discs (CDs) or a user's system. The latter are often referred to as media encryptors.

With the exception of some system files, media encryptors encrypt the entire contents of the drive. Media encryptors need to leave some system files unencrypted so that the computer can boot. Most of these files can have their integrity protected by a cryptographic checksum; this will not prevent a tamper attack, but it will alert the user that that data has been altered. Some system files, however, contain data that changes when the computer is booted. These files cannot be protected. The mechanisms implemented by media encryptors provide:

- Encryption of system files.
- Integrity of the contents of the data storage media.
- Confidentiality of the contents of the data storage media.
- Integrity of the workstation by verifying the Basic Input/Output System (BIOS) and ensuring that configuration and program files are not modified.
- Recovery of data if the original user is no longer able to access the media.
- Key management support, i.e., key generation, distribution, deletion, destruction, and revocation.

File encryptors typically implement a Graphical User Interface (GUI) that allows users to choose files to be encrypted or decrypted. This protects individual files, but it does not protect all of the files on the drive. The mechanisms implemented by file encryptors provide:

- Encryption of selected files.



- Integrity of the contents of the protected file.
- Confidentiality of the contents of the protected file.
- Authentication of a file's source.
- Allow the exchange of encrypted files between computers.
- Recovery of data if the original user is no longer able to access the file.
- Key management support, i.e., key generation, distribution, deletion, destruction, and revocation.

Many applications generate temporary files that may contain user data. These files are normally erased when the application is closed, but when the application does not close in an orderly fashion; these temporary files may remain. Some OSs do not actually erase data when files are deleted. Instead, they alter the name of the file in the file allocation table. The user's data remains on the hard drive until the space is reallocated to another file and overwritten. Thus unencrypted and potentially classified user data can remain on the hard drive after system shutdown, either by failure to erase temporary files, or by design of the OS's erasing function.

There is a wide range of possible architectures for the KMI/PKI needed to support file protection. Possibility range from a user having complete control over key generation and distribution to a hierarchical architecture involving a complex CA. KMI/PKI is discussed in detail in Chapter 8 (Supporting Infrastructures).

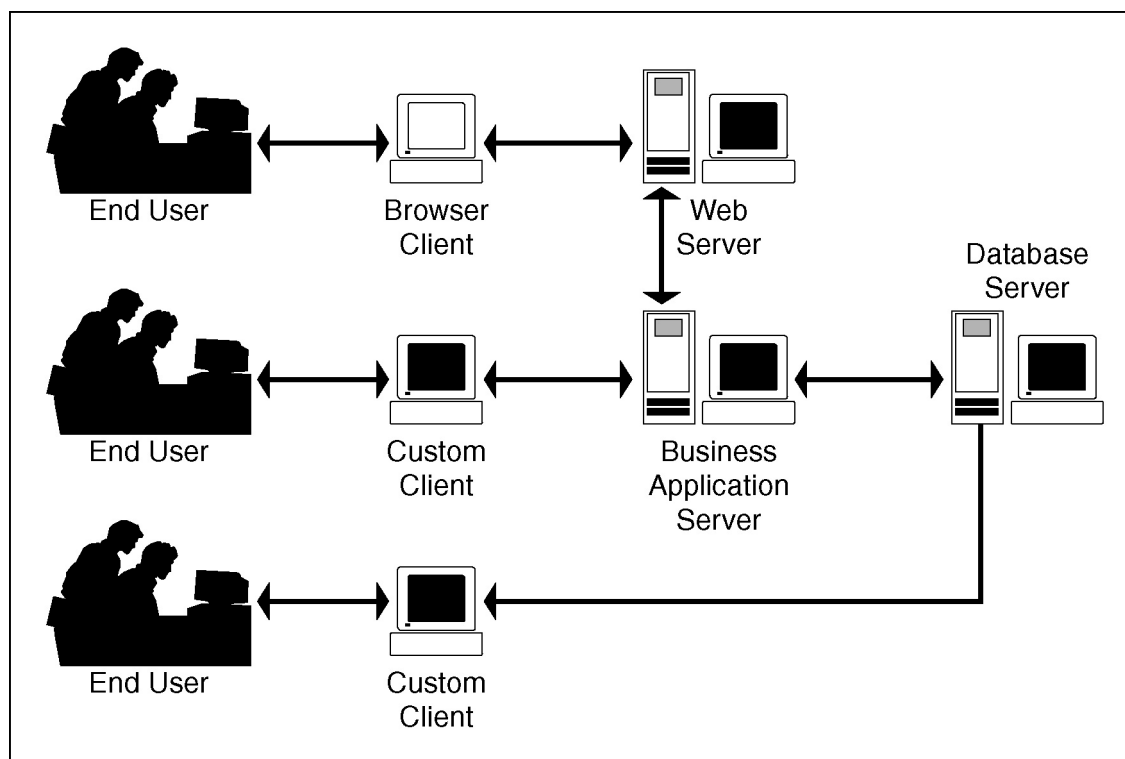
#### 7.1.5.2.4 Mission Specific Applications

Mission specific applications can be as simple as a database making its data available through a fronting web server. They can be as complex as a complete travel service that checks and books airline, hotel, and rental car reservations through a web browser, passes the information to the user via e-mail, and keeps the whole system secure with file encryption. These systems typically rely on existing COTS products, such as web servers and clients and database management systems. As security is only one of many factors involved in the selection of such products, many desirable security features may not be present. In addition, legacy systems with very little security must often be included as part of the solution.

Mission applications need to enforce a definition of security that is specific to the application at hand and to the circumstances of its use. Thus, the security challenge is to combine many less-than-ideal generic component-level security services into a cohesive, meaningful application-level definition. This is a significant information system security-engineering task.

Mission applications are often custom-built using a multi-tiered approach. They are composed of several distinct tiers. The three-tier model typically has a presentation layer, a business process layer, and a database layer. A conventional client-server system uses a two-tier approach. A system can have multiple separate application layers creating multiple tiers. Collectively these systems are referred to as "n"-tier systems. Figure 7.1-1 illustrates the multiple tiers of a system. Different systems will place different numbers of layers between the user and the data. Some systems may simultaneously support multiple access paths to the data. Other systems may enforce a consistent number of tiers for all users.

There are many ways in which a mission application can be secured using readily available technology. Some of these enable the construction of new security-enabled systems. Others allow security to be retrofitted to existing systems and components. In all cases, they are extensions of the security provided by the various system components.



**Figure 7.1-1 Custom N-Tier Application**

### 7.1.5.3 Software Download

Planning for the secure update or download of software must begin early in development and be continued throughout deployment. Three types of software downloads will be considered: firmware updates, software updates, and new software distribution. In all cases, the most critical aspects of software downloads are the integrity of the downloaded software and the authentication of the origin of the software. In some instances, confidentiality of the download may be required. Other aspects that may be important are validity periods, usage limitations, effects of the download on system data, and auditing of the download installation.

#### 7.1.5.3.1 Firmware

The key to managing firmware updates, typified by the recent update of modem software to support a new 56k standard, is planning the hardware support. The hardware must verify the integrity and authenticity of the originator of the firmware updating software. Because firmware is being updated, it can generally (but not exclusively) be assumed that the updated firmware will be processed by the hardware during installation. In general, hardware process-

ing is preferred over software processing because of hardware's better performance and greater resistance to tampering.

Planning for a firmware update must begin with the development of the initial product. Steps that must be taken during initial product development include the following.

- Generate a public/private key.
- Make the public key information readily available.
- Determine whether symmetric or asymmetric cryptography will be used to provide confidentiality of the code which updates the firmware.
- Generate and store symmetric key material if appropriate.
- Field the initial product.

Updating the fielded product requires the following steps to be taken by the firmware developer.

- Generate the code that updates the previously installed firmware.
- Cryptographically hash the updating software.
- Sign the hash with the private key material.
- Encrypt the package (software, hash, and signature).
- Distribute the package.

The deployed system user should then use the product being updated to decrypt, verify signature, verify integrity, and install the update package. Update status, including failures, should be reported through a user interface.

**Integrity** – Package integrity is provided by cryptographically hashing the contents of the package.

**Authenticated Origin** – Signing the hash provides proof of origin. The private aspect of the public/private key pair must be appropriately protected.

**Confidentiality** – Confidentiality can be provided using a symmetric key provided to the user or by using the user's public/private key pair to create a single-use private key for each download. The latter method is more secure, but requires more up-front planning.

**Other Security Services** – Other security services can be provided by hardcoding information in the initial package, or including information for processing in the package. For example, downloading an object that can only be used for a limited time could be governed by validity dates on the signature, coding in the object broker to allow a fixed period of use on each download. In all cases it is important to remember what the security objective is, and manage a chain of trust throughout the achievement of that goal.

### 7.1.5.3.2 Software Update

Software updates are the distribution of modifications made by the developer to software that already resides on a system. This would include service updates to software packages such as Windows NT or Microsoft Office. It would also include distribution of active content code (e.g.,

Java, ActiveX, objects in Distributed Component Object Model (DCOM) or CORBA, macros, etc.). During the download some known trusted piece is already in place to verify the security.

Software updates and active code are managed much like firmware updates. The major exception is that software updates may not be able to rely on hardware storage of key material, so the level of assurance is likely lower than with firmware updates. Also, for most active content, there is a virtual machine of some sort limiting (or at least managing) the operation of the active code, e.g., Java Sandbox, macro interpreter, etc.

### 7.1.5.3.3 New Software Distribution

New software is best distributed on hard to modify media, e.g., CD-ROMS, in tamper resistant packaging with unique vendor identification, e.g., holographic labels. These labels are widely used by commercial vendors to prevent fraud. Some software distribution includes side programs to verify authenticity of the package, or are self-checking. However, since anyone can write code that appears to verify or self-check other code, these mechanisms are not particularly useful.

## 7.1.6 Cases

The potential for insider attacks alone makes it paramount for the security mechanisms described herein to be implemented to some degree for all applications and on all workstations. The degree at which the security mechanisms need to be implemented depends on the potential damage a successful attack could cause. Several “cases” can be defined that are based on the sensitivity (security classification) of the workstation users, associated threat, and enclave configuration. High sensitivity workstations are assumed to employ complementary confidentiality, integrity and availability mechanisms, e.g., strong authentication, encrypting and signing files and e-mail, etcetera. As the sensitivity-classification difference between workstations and individuals in an enclave increase, the need for the countermeasures increases. As the size of the enclave increases the need for coordinating and managing the security within the enclave increases.

### 7.1.6.1 Cases within the Enclave

The following list of cases represents different environments where security mechanisms are needed on the workstation applications to protect information within the enclave boundary.

- 1) Individual user with Unclassified/Sensitive But Unclassified (U/SBU) personally sensitive information within an unclassified enclave.
- 2) Individual user with classified/restricted information within an enclave of equal level/value.
- 3) Subnet of users with U/SBU personally sensitive information within an unclassified enclave.
- 4) Subnet of users with classified/restricted information within an enclave of equal level/value.

## 7.1.6.2 Cases Transiting the Enclave Boundary

Although cases involving information transiting enclave boundaries are handled by other sections of this Framework, the applications can provide another layer of protection for this information. The following list of cases represents different environments where the application can provide this additional layer of protection.

- 1) Individual user with U/SBU personally sensitive information communicating with unclassified network, e.g., Internet.
- 2) Individual user with classified/restricted information connecting to network equal level/value.
- 3) Remote SBU user connecting through unclassified network to SBU Local Area Network (LAN) (Remote Access).
- 4) Remote classified user connecting through lower level network to classified network (several subcases by deltas in levels) (Remote Access).
- 5) Unclassified/SBU/Restricted but “lower value” information LAN to large, open unclassified network t, e.g., Internet (many adversaries of varying capabilities).
- 6) SBU or Classified (valuable information) LAN to a network of the same classification/value (less open).
- 7) Classified LAN or LAN containing valuable information communicating through a lower level network to another network of equal classification/value (System High Interconnects).
- 8) Classified LAN or LAN containing (highly) valuable information communicating with a lower classification/value network (High-to-Low, Multi-Level Security (MLS)) (multiple sub-cases exist for varying deltas between information on the LAN versus the Wide Area Network (WAN)).
- 9) Classified LAN or LAN containing valuable information connecting to same classification/value/organizational WAN, which has limited connections to lower classification/value/external network, e.g., Secret LAN connected to Secret WAN that is also connected to an Unclassified WAN.
- 10) Sensitive, restricted, or compartmented information LAN/subnet to corporate net/Intranet.

The first four cases involve a single individual workstation connecting to a similar security level component, employing a potentially lower sensitivity level transmission media. Cases 5 through 7 involve interconnected networks at essentially the same sensitivity level, employing unprotected (lower sensitivity level) transmission media. Cases 8, 9, and potentially 10 involve high-to-low connections (which may jeopardize interconnected high level systems that are not aware of the low connection). Case 10 may involve a range of differences in information value of the subnet versus the network.

## 7.1.7 Framework Guidance

This Framework provides a characterization of the security features and assurances needed to provide information assurance in today's highly networked and richly interconnected environments. It is the applications that process and circulate information. Providing affordable security-enabled applications is paramount to providing overall information assurance. If implementing security-enabled applications involves a significant financial investment, organizations and users will be reluctant to implement them. Application developers must strive to develop security-enabled applications that meet the user needs without adding extras that drive the costs to prohibitively high levels.

This section will not provide guidance for each of the cases presented in Section 7.1.6 (Cases). Instead, the section will provide guidance that can apply in all cases. The specific requirements for each case and type of application will be provided in the form of protection profiles that support the DoD defense-in-depth strategy.

### 7.1.7.1 User Interface

A security mechanism that is cumbersome to use will not be used. The importance of an intuitive and burden free user interface can not be overemphasized. This easy to use user interface is necessary for day-to-day operations. The user interface is also involved in key management, at least during the initial start-up period. It is important that the user's involvement in key management, both procedural and electronic, does not cause undue burden. If this key management is a burden for the user, encryption and digital signatures will not be widely accepted or used within the organization. The user interface should keep the user apprised of security-related events and pertinent information. Following is a list of these events and pertinent information.

- Outgoing information has been encrypted and/or digitally signed.
- Incoming information is encrypted and/or digitally signed.
- The identity of the person who encrypted and/or digitally signed the incoming information.

### 7.1.7.2 Security Mechanisms

Not every vendor implements security mechanisms in the same way. Providing configurable mechanism options increases the chance that products from different vendors will operate together. These mechanism options can include the algorithms and associated key lengths supported by the application and the protocols used to transfer information between users, e.g., S/MIME or MSP for messaging. A tradeoff must be made between the need for the secure application to be able to support a number of mechanism options and the need for the application to be inexpensive and easy to use. Generic applications should have a means of determining the common mechanism options that are implemented when two or more applications attempt to interoperate.

There are two ways to add security mechanisms to applications. First software plug-ins, e.g., the LJI Armor e-mail products, can be added to existing non-secure applications to add

security features. Alternatively, security mechanisms can be directly integrated into the application. That is the security mechanisms are added during product development and are an integral part of the application, e.g., the RSA and FORTEZZA® modules integrated into the Netscape web browser. Although there are advantages to both methods, integrating security mechanisms directly into applications is preferable. Security should be an integral part of an application not an after thought. The following is a list of constraints that security-enabled applications should meet.

- Applications with similar functions interoperate, e.g., secure e-mail packages can communicate with different secure e-mail packages.
- The user has the choice to enable security mechanisms selectively for each message/file being sent.
- The user should be able to apply encryption only, digital signatures only, or both encryption and digital signature to information.

The encryption and digital signature mechanisms, e.g., algorithms, key lengths, random number generators, etc., should be of sufficient strength and in keeping with the current legal policies for the environment in which they will be used.

### 7.1.7.3 Compromised Keys

A policy is needed to address compromised keys. There are a number of issues surrounding such a policy. These issues include who determines a key has been compromised; where the list of compromised keys is maintained; how this compromised key list is disseminated; and what actions need to be taken once it is discovered that a received certificate has been compromised. Electronic mechanisms must be in place to enforce the set compromise key policy. The security administrator should be able to configure the compromise key enforcement mechanisms, as appropriate, in order to implement the site's policy.

Determining that a key has been compromised and thus is to be put on a compromised key list would normally be done in two ways. The first would be the case where the individual originally issued the key no longer is connected with the issuing organization. Perhaps the best example of this would be an individual losing their access to classified information upon leaving Government service. Leaving a private company and thus access to the company's important confidential information, however, could be just as critical. In this case, the maintainer of the compromised key lists simply adds the departing individual's key to the list. The second method entails each key having a validity period. If once the key has exceeded its validity period and no action has been taken to renew the key either by the issuer or individual initially issued the key, the key is considered compromised. The key is then added to the compromised key list. A combination of both is required.

It is important for the compromised key list to be maintained in a central location that is easily accessible by all the users. The policy must establish the identity of the trusted centrally located server or servers. The policy must identify under what circumstances the users must check with the server maintaining the compromised key list.

There are two ways a compromised key list can be disseminated, pushing or pulling. The entity maintaining the list of compromised keys could push the list out to the user, or the users could pull the information off the list maintained at one or more central sites. Pushing the list out to

the users should be done only in extreme cases, e.g., continued use of a compromised key could cause serious damage to the system's security, because it requires the server to establish a trusted link with all users, and causes a bottleneck in the system. The primary method used to disseminate a compromised key list should be the user pulling the list from the trusted centrally located server or servers. The policy should set a specific period for which the key list is valid. After that period has expired, the user is required to pull a new list.

When it is discovered that a certificate has been put on the compromised list the policy must identify whether the choice of the action to be taken should be automatically implemented in the code or determined by user. The code must be able to support the following possible actions.

- Compromised key is marked as compromised, the system maintains the original public key information so the user is still free to decrypt messages and verify signatures.
- Compromised key is replaced with a compromised key certificate that notifies the user when a message using the original compromised key is received but does not decrypt messages or verify signatures.
- Compromised key is deleted.

### 7.1.7.4 Password Practices

A policy must be set for the site designating good password usage practices. The FIPS Publication 112-1 "Passwords Usage" provides information on good password practices. For example, good practices include a minimum password length of 8 alphanumeric characters, a maximum period of password usage, and invulnerable to simple dictionary attacks. Electronic mechanisms must be in place to enforce the usage of good password practices. This is particularly important when the passwords are being used to protect the private key information. The security administrator should be able to configure the password enforcement mechanisms as appropriate in order to implement the site's password policy.

### 7.1.7.5 Technology Gaps

The tools, mechanisms, and security services necessary for building secure applications are generally available. There are however serious gaps. The gaps come mainly in the difference between known capabilities and needs, and available solutions. Finding a full vertical solution slice is quite difficult. Such a solution would include tokens, certificate infrastructure, and applications that all understand each other, use the same version of certificate with the same fields, and all use the same (where appropriate) standards for interoperability.

This ideal solution is nearly impossible to find today. There is market fragmentation at virtually every horizontal level. Tool vendors use different algorithms, service vendors use different protocols, standards are not completely defined for interoperability, the certificate infrastructure uses different certificate extensions (sometimes with different meaning or intent), different directory services and query modes exist, and the applications use differ-



ent standards, or different protocols. Examples of this fragmentation abound. E-mail is an excellent example: the Defense Message Service uses MSP mail formats, the commercial world uses S/MIME and OpenPGP. Some applications use X.509 version3 certificates, others still use X.509 version1.

This gap in vertical solutions is expected to be resolved as products from larger vendors (Sun, Microsoft, Lotus, and IBM) begin to appear on the market. In the meantime, vertical solutions are often proprietary, and thus of limited interest to the Government. As the gap in basic solutions disappears, the technology gaps will become more focused on the capability and security provided by the products, with some implementations simply being more robust than others. Additionally, new technologies (smart cards, PCMCIA cards, dynamic HyperText Markup Language (HTML), Virtual Reality Modeling Language (VRML) and others) will need security, yet find it lacking the first generation of products.

Testing, evaluation, and use of various products will eventually highlight where the real security gaps in applications are, and what can best be done about them.

## References

1. Emmett Paige, Jr., "Selection of Migration Systems," Assistant Secretary of Defense Memorandum, November 1993.
2. Linn, J., "Generic Security Service Application Program Interface," RFC 1508, November 1993.
3. Adams, C., "Independent Data Unit Protection Generic Security Service Application Program Interface (IDUP-GSS-API) ," Internet draft 4, February 1996.
4. X/Open, "X/Open Preliminary Specification: Generic Cryptographic Service API," draft 8, April 20, 1996.
5. Kaliski, B., "Cryptoki: A Cryptographic Token Interface, Version 1.0." RSA Laboratories, April 28, 1995.
6. Linn, J., "Generic Security Service Application Program Interface," Version 2, Internet draft 5, February 1996.
7. National Security Telecommunications and Information Systems Security Committee, "National Information Systems Security Glossary," NSTISSI No. 4009, 5 June 1992.
8. National Computer Security Center, "An Introduction to Certification and Accreditation," NCSC-TG-029, January 1994.
9. National Computer Security Center, "A Guide to Understanding Security Modeling in Trusted Systems, NCSC-TG-010," October 1992.
10. Microsoft Corporation, "Application Programmer's Guide: Microsoft CryptoAPI," preliminary version 0.9, January 17, 1996.
11. NSA Cross-Organization Team, "Security Service API: Cryptographic API Recommendation," First Edition, National Security Agency, June 12, 1995.
12. Schneier, Bruce, Applied Cryptography, 2<sup>nd</sup> Edition, John Wiley & Sons, 1996.

**This page intentionally left blank.**

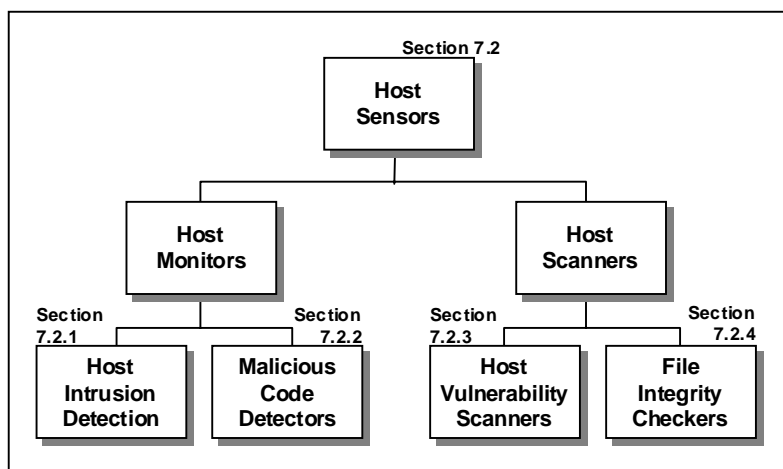
## 7.2 Host-Based Detect & Respond Capabilities Within Computing Environments

A fundamental tenet of the defense in depth strategy is to prevent cyber attacks from penetrating networks and compromising the confidentiality, integrity, and availability of enclave information, and to detect and respond effectively to mitigate the effects of attacks that do. The host-computing environment is the final line of defense within that overall defense in depth strategy. These workstations and servers can be vulnerable to attacks (via poor security postures, mis-configurations and software flaws) and must be factored into the computing environment's protection approach.

Previous discussions have addressed enclave boundary protection and associated detect/respond approaches. While these technologies offer perimeter and access controls, "authorized" internal and remote users can attempt probing, misuse, and malicious activities within an enclave, particularly when they have become authenticated in some manner (either as an authorized user or by impersonating an authorized user) to a host computer.

Detect and respond capabilities are complex structures that run the gamut of intrusion and attack detection, characterization, and response. The various Detect aspects of Detect and Respond are actually measurement services. Intrusion detection, network scanning, host scanning, and the like are measurement functions that, on a continuous and/or periodic basis, determine the effectiveness of the deployed protection systems. In themselves, Detect capabilities are not Protection measures. The Respond aspect can initiate changes to existing protection systems (e.g., automatic disabling of a user's account after too many failed login attempts) or the deployment of additional protection measures (e.g., implementation of strong authentication systems). The local computing environment is the logical location for host-based sensors (within an enclave environment). This section addresses host-

based sensors including those that operate in near real time and those that operate off-line. Specific host-based sensor technologies addressed in the Framework are shown in Figure 7.2-1. Sections 6.4 (Network Monitoring Within Enclave Boundaries and External Connections) and 6.5 (Network Scanners Within Enclave Boundaries) of the Framework provide similar guidance for network sensor technologies. There is commonality in a number of the topics covered in the respective sections of the two chapters. Rather than cross-referencing across the sections, each is structured as stand-alone for the convenience of the reader.



**Figure 7.2-1 Breakdown of Host Sensor Technologies**

There are a number of functions (e.g., intrusion characterization and response formulation) typically performed by analysts that use the information provided by locally deployed sensors. Local environments have the option to implement as much or as little above the sensors as they feel is prudent, obtaining services and support from the system infrastructure as necessary. Section 8.2 (Detect and Respond as a Supporting Element) of the Framework provides an in-depth discussion of the various Detect and Respond processes and functions in the context of a supporting Information Assurance (IA) infrastructure capability. It also offers guidance on technologies for processes beyond the sensors, but recognizes that they can be implemented at any level in an enterprise hierarchy (including a local computing environment).

Host-based sensors covered in this section include host monitors (intrusion detection and malicious code detector technologies) and host scanners (host vulnerability scanners and technologies for software integrity checking). The section provides an overview of each relevant technology, general considerations for their use, rationale for selecting available features, deployment considerations, and a perspective on how these technologies are typically bundled into products. The section concludes with sources for additional information and a list of references that were used in developing this guidance.

## 7.2.1 Host Monitors—Intrusion Detection

Today, most operating systems and applications generate some form of audit trail. Originally, it was intended that a security administrator would review the audit logs looking for suspicious events. While currently in practice, such manual processes do not scale with the limited personnel resources typically required to review such logs. Many enterprises do not make use of audit logs (or the tools used to facilitate their analysis) for two major reasons. The tools themselves are heavily dependent upon the user's ability to understand the types of attacks and vulnerabilities. And as the number of users, operating systems, applications, and databases grow, so do the audit trail file sizes, which often consume too much storage and could result in denial of service problems. Often, information technology (IT) operations staff are forced into deleting or disabling audit trails in order to avoid costly disruptions to their networks and information processing systems.

### 7.2.1.1 Technology Overview

The goal of a host intrusion detection system (IDS) is to identify unauthorized use, misuse, and abuse of computer systems by internal network users in "near real time." As discussed in Section 6.4 (Network Monitoring Within Enclave Boundaries and External Connections), similar structures and technologies are also available for performing comparable functions using network-based information.

Host IDSs are based on the principle that an attack on a computer system will be noticeably different from normal system activity. An intruder to a system (possibly masquerading as a legitimate user) is very likely to exhibit a pattern of behavior different from the normal behavior of a legitimate user. The job of the IDS is to detect those abnormal patterns by analyzing numerous sources of information that are provided by the existing systems. Two major techniques used to detect intrusions are statistical analysis and rule-based expert system analysis, as discussed below.

- *Statistical Method* attempts to define normal (expected) behavior, while the expert system defines proper behavior. A popular way to monitor statistical measures is to keep profiles of legitimate user activities. These profiles may include such items as login times, CPU usage, favorite editor and compiler, disk usage, number of printed pages per session, session length, error rate, etc. The IDS will then use these profiles to compare activity with past user activity.
- *Expert Systems* are another means of detecting possible attacks on a computer system by searching for breaches of policy. They typically use a rule-base expert system to analyze the audit trail records, and will try to discover attacks based on the information contained in the rule base. The expert system is able to pose sophisticated queries to the rule base to answer conditional questions based on sets of events. The main problem with this method is determining exactly what the rules should be and what kinds of attacks can be detected using this method.

## Detection Approaches

Anomaly and Misuse Detection—This category of technologies attempts to identify benign or intentional unauthorized system use, applying anomaly and misuse detection technologies to detect changes in the patterns of utilization or behavior of the system.

- *Anomaly Detection* techniques assume all intrusive activities deviate from the norm. These tools typically establish a normal activity profile (a statistical model that contains metrics derived from system operation) and then maintain a current activity profile of a system. Observed metrics that have a significant statistical deviation from the model are flagged as intrusive. When the two profiles vary by statistically significant amounts, an intrusion attempt is assumed.
- *Misuse Detection* systems attempt to identify authorized users' misuse of computing resources. Detection is performed by looking for the exploitation of known weak points in the system, which can be described by a specific pattern or sequence of events or data (the "signature" of the intrusion). Such activity may include visiting unauthorized Internet sites, navigating around a system to areas that have been explicitly identified as "off limits," or using an application for activity unrelated to work. Misuse detection systems typically rely on an administrator defining activity that is considered misuse through the use of configuration files. The information in the configuration files can then be compared with the activity that occurs on the system; misuse is assumed when there is a match between the two.

## IDS Tuning Options

Typically, a host-based IDS provides capabilities for tuning its operation to a particular host and enterprise environment. Depending upon a particular IDS implementation, it is often possible to pre-determine the types and specific attacks to be monitored, what the response will be for each detected intrusion (e.g., generate an alarm, record, take a mitigation action, etc.), and characterize the class (e.g., indication of the importance or severity) of each alarm generated. This can be driven both by anticipated authorized activity on the host as well as the general information system usage characteristics across the enterprise. In this way, it is possible to focus the host IDS on specific events of interest, dependent upon what threats have been identified as relevant to the particular host environment, and the response the IDS will

have upon the detection of events. An IDS should not be deployed without a concept of operations and a set of well-defined goals, responses and tuning approaches in mind.

Often, tuning involves evaluating the operation of the IDS at initial activation for a period of time (some implementations do “self-tuning”) and then “tuning out,” or de-sensitizing, the monitor where appropriate. In some cases, an increase of sensitivity may be required, but most technologies come highly sensitive “out of the box.” Anomaly Detection elements usually have some learning curve to determine normal patterns and distributions of activity. Finally, as mentioned above, other adjustments can be made to deselect some activities and to add in others based upon the analysis and correlation of alarms and alerts with other measures in the system.

## Response Options

While the sensors detect and collect information about intrusions, it is the analyst that interprets the results. Host-based IDS agents watch aspects of a host or server security such as operating system log files, access log files, application log files, as well as user-defined application policies. If a breach of policy occurs, the host IDS can react by logging the action, alerting the administrator (notify a console, send e-mail, beep a pager), disabling an account, terminate an intruder’s session, shut the system down, or execute a command, in some cases stopping the action prior to execution.

## Reporting Mechanisms

When the host IDS determines that the criteria have been met for declaring an intrusion, anomaly, or misuse event, it is generally configured to either signal alerts to a console interface or a centralized management station where information can be recorded and brought to the attention of an administrator. Some host IDSs can send e-mails (from the central console or individual agents) to alert an operator to events or initiate telephone pages if properly provisioned and configured.

As with network IDSs, many host IDS central-reporting consoles or systems come with database components that allow the general manipulation or correlation of event data, as well as the generation of a wide variety and format of reports, both graphical and numerical.

### 7.2.1.2 General Considerations for Selecting the Technology

A host-based IDS, rather than scanning network packets, watches the audit logs or other system resource events and activities on each monitored computer for signs of misuse. The benefits of host-based IDSs are that they are easy to configure for individual servers and applications. They provide tailored security since they can monitor specific operating system (OS) or application events and can enforce enterprise policies. Only host-based IDSs can detect an intrusion that occurs through the locally attached console. And only a host-based IDS can enforce a user-based reaction policy (e.g., disable the user account, terminate user process) when an attack is detected.

A host IDS is well suited for monitoring specific user and file activity. However, since it cannot detect network-based threats, host-based IDS should be considered as a means to complement

network-based IDSs, supplementing detection of intrusions that may appear to be part of “authorized” traffic flows or otherwise missed within switched environments. While use of both technologies is preferred in general, there are situations where it may be appropriate to use host-based IDS only:

- Network bandwidth is too high to enable network monitoring, or too low to justify the expense of a network IDS.
- The network environment is highly switched (logically segmented), without span ports on the switches, or the mesh is too large, making the number of sensors prohibitive.
- The topology is too highly distributed (either geographically or logically segmented).
- Organizational/domain communities of interest or ownership issues (e.g., different organizations own the network versus the hosts or some subset of hosts—and they don't communicate well).
- Privacy/Consent issues, as it is much easier to have a “consent to monitor” policy when logging into a host than a network.

A classic example where host-based IDSs are the only practical approach is a high performance computing community, where a loose coalition of high-end computing environments that share data, but the owners of the processing capacity do not own the network over which they communicate.

Host based IDS performance varies by the number of standard attack definitions and enterprise-specific policies being monitored, as well as the number and type (compute-bound versus input/output-bound) of processes executing on the given host. In addition to the speed of the host and its components, another factor is the enterprise architecture for host management.

Although intrusion detection and response systems are important components of an overall enterprise security program, devices currently in use have many flaws. Host-based IDSs rely on inherently after-the-fact analysis of audit data to detect suspicious activity and anomalies, and they are difficult to scale for use in large enterprise environments. A concern with a host IDS is the potential computational overhead imposed on mission-critical servers and hosts whose security is being monitored (since the IDS resides on the same machine). Another consideration is complexity of deployment and administration, which varies by the number and types of different servers being protected. A host-based IDS cannot address attacks that exploit protocol vulnerabilities, and, since they analyze data from the audit trails, reaction to an attempted intrusion is typically not in real time. Another limitation of host-based IDS is that the access to audit trails is available only at the OS level or at the application level. This is why host-based IDSs should be implemented in the context of a total defense in depth security posture with a comprehensive approach to enclave boundary security.

Table 7.2-1 below provides a synopsis of advantages and disadvantages associated with host-based IDS technologies.



**Table 7.2-1 Host-Based IDS Considerations**

Advantages	Disadvantages
<p>Provides real-time measure of the adequacy of a system's access control and protection measures.</p> <p>Systems can monitor information in terms of who accessed the system.</p> <p>Systems can map problem activities to a specific user identification.</p> <p>System can track behavioral changes associated with information system misuse, which is typically an "insider" of the information system.</p> <p>System can operate within an encrypted environment.</p> <p>System can operate in switched network environment.</p> <p>Systems can distribute the load associated with monitoring across available hosts on large networks.</p>	<p>Network activity is not visible to host-based sensors.</p> <p>False alarm rates are high with current technologies</p> <p>Activating audit mechanisms can incur additional resource overhead on the system.</p> <p>Audit trails used as data resources can take up significant storage space.</p> <p>Operating system vulnerabilities can undermine the integrity of host-based sensors and analyzers.</p> <p>Management and deployment costs associated with host-based systems are greater than in other approaches to implementing an intrusion detection system.</p> <p>Host-based sensors are more platform-specific, which adds to their cost and required operator expertise.</p>

Finally, the degree to which the host-based IDS is configured to monitor a particular system should be dependent upon the sensitivity of the information being processed by that system or the criticality of the system to the integrity and availability of the entire enterprise. Host-based IDS systems come with operational and managerial burdens. These include alerts that require very specific administrator intervention for examination, implementations that may be available only for specific operating systems, and system performance impacts on the host. A broad deployment of host-based IDSs requires careful planning, or its use is not recommended. A threat/risk assessment is strongly recommended to identify particular hosts on which to add host-based IDSs. Then a careful deployment activity and continual monitoring for performance impact or operational degradation should be undertaken.

### 7.2.1.3 Important Features

When considering the selection of a host-based IDS, there are a number of features that should be considered. This section identifies important features for selection. The section that follows discussed rationale for the selection of these features.

#### Detection

- Support for service startup detection.
- Ability to detect registry changes.
- File-watching capability.
- Ability to profile normal activities and detect variations from the norm.

#### Signatures

- The number of events/signatures.
- File or message integrity checking based upon cryptographic algorithms, not simple checksums.

- Customizable system checks.

## Operations

- The actual deployment and management capabilities of the complete IDS system (e.g., number of agents that can be connected to a single manager, and number of managers that can report to a single console).
- Ability of the auditing process to automatically reset itself.
- Support for remote management.
- Ability to integrate with network-based modules; how well the tool works in a heterogeneous environment becomes a critical factor when considering true enterprise class IDS tools.
- Survivability characteristics (self-recovering from power loss, resource failure, component failure and similar situations).

## Response Options

- Configurable, automated, rule-based response capabilities.
- Account blocking, access control changes, etc.
- Ability to coordinate responses across multiple host platforms (e.g., disable the same account on **all** enterprise systems or at the NT Domain Controller).
- Integrated response with network-based tools.

## Reporting Options

- Capability to perform Simple Network Management Protocol (SNMP or trap) alerting to a centralized system management operation.
- Capability to notify personnel using e-mail alerts and a variety of other contact measures (pager, fax, etc.).
- Capability to automatically execute programmed scripts on alerts at management system or console (also partially a response function).
- Ability to generate customized reports on an as needed basis.
- Ability to capture events into a standardized database system.

## Performance

- Balance between the overhead required to audit OS and application activity logs and the ability to react to infractions.
- Data log impact upon system resources (since host-based IDS generate log files as well).

## Platform

- The specific types of platforms (e.g., operating systems) upon which the tool operates.
- Minimum Platform Configuration.

- Memory Requirements.
- Disk resource requirements.
- Ability to handle crossover when reporting between platforms.

## Console Considerations

- Operator Interface—Type of command and monitoring provisions available to an operator.
- Mark As Analyzed—Ability to clear or mark selected alarms that have been reviewed.
- Drill Down—Ability to provide additional information for selected events.
- Correlation—Tools to correlate events based on source, destination, and type.
- Report Generation—Ability to generate reports upon event detection and as periodic summary reports.
- Integrated Industry-standard database.

### 7.2.1.4 Rationale for Selecting Features

Detect and respond capabilities exemplify the necessity of integrating operations and personnel considerations with the selection of technology solutions, consistent with the overall defense in depth philosophy. As indicated earlier, host-based monitoring does not itself offer protection from intrusions or attacks. It should be considered more as instrumentation that monitors (and “measures”) the effectiveness of a host computer’s existing protection structures. It is up to system administrators (support and operations staff) to interpret the outputs and reporting of the IDS and initiate an appropriate response. If full time operators<sup>1</sup> are not available to interpret and formulate responses based on the IDS outputs, then IDS implementations will not typically add real value. In this case, it is likely that IDS deployments should not be considered. Otherwise, when considering features for an IDS, there are a number of factors to be considered that are based on how the IDS is intended to be used, whether full or part-time operators will be available, and the skills of the operators to interpret the results.

## Detection

Within the host-based IDS environment, most detection technologies actually use a mix of both signature matching and anomaly/misuse detection. There are advantages to both approaches. While signature-based IDSs are the traditional method, they typically lack the ability to detect new (or modified) versions of attack patterns. While many intrusions (typical of novices) use standard attack sequences (often downloaded from hacker bulleting boards), an accomplished adversary will have the capability to create new attacks, or modify old attacks and thus thwart traditional signature detection mechanisms. Anomaly and misuse detection approaches (e.g., statistical profiling, unauthorized system resource use or modification monitoring) have greater

---

<sup>1</sup> Ideally operators should be available on a 24x7 basis. The number of operators will depend on the traffic loads and anticipated numbers of incidents. It is not uncommon to experience hundreds of thousands of intrusion alerts per day, and each has to be investigated to determine which, if any are serious threats.

flexibility for identifying new or modified attacks (since they monitor network usage or behavior). These are also the only mechanisms currently available to monitor actions of otherwise authorized users for inadvertent or intentional misuse. They can sometimes be more complex to operate and manage, but in most technologies, the degree to which each aspect (signature vs. misuse/anomaly) is enabled in most technologies is configurable.

As always, any decision is based on level of risk, anticipated performance, cost (for purchase, deployment and operation), and operational impact. The Framework recommends consideration for deployment of multiple attack detection schemes. In this way, there is a greater likelihood of detection by at least one of the mechanisms deployed.

## Signatures

Concerning a signature-based IDS or component of an IDS, it is desirable to have as many signatures as possible used for detection. However, there is usually an inverse relationship between the number of signatures, and the response time for possible detection, and the amount of traffic that can be monitored. Since the “lists” of possible attacks change frequently, it is strongly recommended that the IDS be capable of dynamically loading signatures. It is usually operationally more feasible and efficient if the downloading is handled on an enterprise (or at least site) basis. Most vendors that offer dynamic loading of signatures provide periodic updates to their signature base. While the update periods differ among vendors, a good rule of thumb is that more is better. If operators have the skills to create custom signatures, then having the ability to support user-defined attacks is also desirable, particularly if custom attacks are found in one of your sites.

## Operations

It is desirable for the IDS to be easily configurable according to the security policies of the information system that is being monitored. Consideration should also be given to its ability to adapt to changes in system and user behavior over time (e.g., new applications being installed, users changing from one activity to another, or new resources being available that cause changes in system resource usage patterns).

By its nature, IDS sensors are located where intrusions are anticipated. Thus, it is important that an adversary not be capable of modifying the IDS to render it ineffective. It is desirable that the IDS component within a host system environment is able to perform self-monitoring, detect unauthorized modifications, and notify an attendant console. To simplify recovery of operations after an intrusion, it is also desirable that the IDS be able to recover from system crashes, either accidental or caused by malicious activity, and upon startup, be able to recover its previous state and resume its operation unaffected.

## Response Options

Many available solutions offer automated response options that seem on the surface to be very desirable. They infer that little or no human interaction is involved, as the devices can provide an immediate response. There is a serious pitfall to consider, however, before these options are deployed. First, it is not uncommon for a host IDS to identify as potential misuse many of events daily, depending on where it is employed, that are in fact characteristics of the “normal” host usage load. Often, without careful tuning, the number of false positives may be high,

giving frequent rise to unwarranted indications of intrusions. Automated responses that terminate user sessions, modify access controls, throttle processes, or actual shut down a system can often cause severe denial of service threats to the network. It is strongly recommended that automated options not be used unless there is some mechanism to control the potential for denial of service.

## Reporting Options

Most host-based IDSs report alarms to an operator console (please refer to the discussion of console features below). The desirability of the level and frequency of reporting is based primarily on the availability and skills of the operators. Some host IDS technologies offer the option of paging or sending e-mail messages to notify personnel of alarms. While these sound desirable, they have the potential to give rise to operational issues. With an IDS detecting thousands of alarms a day, these features have the potential for overloading e-mail servers (creating a denial of service threat themselves), or paging operators extremely frequently at all times of the day and night. Most often, these features are not recommended, at least not until some consistent baseline of normal behavior is identified.

## Performance

Host IDS performance varies based on the available resources (processor, bus architecture, disk space) of the host system, the actual operational applications it is executing, the number and type of processes it experiences during operations, the number of attack signatures employed, and the level and complexity of audit or analysis the IDS is configured to undertake. Unlike network-based intrusion detection sensors, where performance degradation results in the lost of intrusion detection capabilities but not network performance, host-based sensor software can impact the host system itself overall. In each case, a tradeoff must be determined between the level of audit the sensor software is configured to undertake versus the impact on overall system performance. In the case of where existing host performance is already marginal, redesign of the system and sensor software deployment approaches should be considered. Host based IDSs must be deployed very carefully.

## Platform

A major issue for the selection of a host-based IDS is the type of computer skills (e.g., UNIX, NT) required for operators. They will likely need these skills to perform installation, configuration, adjustment, and maintenance. Since a host-based IDS is usually deployed into existing system, knowledge of what is already running on the system and the resources it requires is critical. In addition, the console platform will have to be acquired and maintained, so it may be useful to select a technology that functions on the types of platforms used within the enterprise.

## Console Considerations

As discussed in Section 8.2 (Detect and Respond as a Supporting Element) of the Framework, the primary function of the console is to serve as an aid in the characterization and analysis of the many alarms that will be identified. Operators will have to identify alarms that resulted from authorized use (e.g., false alarms), those that do not offer serious risks to the network, and those that do, as well as gain an initial perspective on the source and impact of possible attacks

**Operator Interface**—The type of interface that is operationally desired tends to be driven directly by operator preference. Novices typically prefer a graphical user interface (GUI) with intuitive operations, pull down screens, and substantial aids available. More skilled operators may prefer command string operations, tailored screen options, and more options for operator customization. It is best if operators can get a hands-on trial evaluation of the console capabilities prior to final selection.

**Mark as Analyzed**—Operators will *typically* be faced with large numbers of alarms that have to be analyzed and cleared. A capability that is usually critical is the ability to selectively keep track of alarms that have been reviewed.

**Drill Down**—Many host IDS consoles display a high level characterization of events in order to display the large number of alarms that are detected. Operators will usually have to access additional details about each alarm to be able to characterize it properly. It is very desirable for the console to be able to provide the additional levels of information when requested by the operator. As with the operator interface, the types of information desired will typically depend on the skills of the operators.

**Correlation**—In the same vein as drill down features, operators will require tools for correlating (e.g., based on source, destination, type alarms and events) in order to identify and properly characterize intrusions and attacks. This is particularly the case in situations where the incidents are distributed in time or location. The ability of the console to integrate the reporting of various host-based and with network-based IDSs and other relevant events is a strong plus, if the operators will use the additional information. Again, as with the operator interface, the types of tools desired will typically depend on the skills and mission of the operators.

**Reporting**—The type of reporting options will depend predominantly on the type of information operators will want to perform their characterization, and the organization's need for reporting to higher levels (e.g., periodic summary reports). It is always desirable for the console to be capable of generating reports that can be created and disseminated with little extra effort beyond the hour-to-hour and minute-to-minute responsibilities that the operators will otherwise have.

## 7.2.1.5 Considerations for Deployment

Host-based IDSs are designed to monitor a single host on which it (or an agent of it) resides. Typically this kind of IDS is able to watch data available from higher levels of protocol stacks, which restricts its ability to monitor activities to those audit trails made by OS or applications. It also can detect the activities that occur locally on the monitored host (e.g., file permission modification and user account setup).

Host-based IDSs fall into two basic configurations: single system and agent/manager. A single system IDS protects one machine by detecting intrusions in the machine's audit logs and other methodologies. A manager/agent host-based IDS application involves placing agents on one, some, or all hosts. In this approach, IDS agents reside on systems that are to be monitored. These host-based systems also rely on the analysis of the operating system event logs and audit processes (among other techniques described above) to detect suspicious activity. Here, these host-based agents are part of a distributed architecture in which the system agents report to a centralized management station. These agents are connected to managers, which in turn are

connected to a central management console. Agents can remotely install/upgrade new versions and attack signature rules. This type of configuration allows security administrators to define and distribute rules from one central location. Some host monitors can also track audit trails from other applications, like firewalls, web servers, and routers. These fall into the category of network-based monitoring capabilities, and are discussed in Section 6.4 (Network Monitoring Within Enclave Boundaries and External Connections) of the Information Assurance Technical Framework (IATF). While the Framework focuses on the technology aspects of an overall IA solution, it is important to note that the value of an IDS is realized only when a competent operator or analyst can interpret the result. This requires operator training to ensure they have the analytical skills and proficiency with tools to make correct interpretations efficiently. It is also very important to develop procedures (e.g., courses of action, standard operating procedures) for operators to follow for all contingencies, but particularly for when serious alarms are discovered.

### 7.2.1.6 Considerations for Operation

As discussed above, most IDS technologies provide the capability to tune the sensor to improve its performance for specific deployments. When an IDS is first deployed, it is prudent to operate the technology for some period depending on the complexity of the deployment to complete this tuning. This provides an opportunity for determining that the IDS is capable of monitoring applications and detecting alarms, and for increasing or decreasing sensitivities. Also, anomaly detection elements usually have some learning curve to establish a baseline for normal patterns and distributions of activity. It also enables other adjustments to deselect some activities and to add in others based upon an analysis of alarms that are triggered.

Tuning enables the IDS to preclude the detection of authorized traffic patterns that may otherwise cause false positive alarm indications. There are two fundamental approaches for tuning. The first is to have knowledge a priori of the usage patterns that could trigger false alarms. The IDS can then be configured (tuned) to preclude these from causing an alarm.

While it is desirable to have the information in advance, it is often not available. The other approach is to run the IDS and have it find incidents and conditions that generate alarms. As alarms are detected, an analyst determines if an actual intrusion was discovered, or if the alarm was the result of a false positive based on normal operation. As indicated above, the IDS can then be tuned to preclude those events from triggering an alarm indication. This form of “discovery” also provides operators an opportunity to become familiar with the technology before it goes on-line operationally.

Tuning should not be thought of as strictly an installation process. This should be done on a regular basis, to refine and focus the detection mechanisms on real intrusions, and reduce false positives throughout its operation.

Once an IDS is deployed, it is recommended that testing of the IDS be conducted to ensure that it is configured correctly and functioning properly. While it is also possible to construct exercises to determine the proficiency of the operators and analysts, it is likely that normal day-to-day operations will provide more than sufficient numbers of real alarms to provide opportunities to assess their capabilities.

## 7.2.2 Host Monitors— Malicious Code or Virus Detectors

Over the past decade, computer viruses<sup>2</sup> have gone from an academic curiosity to a persistent, worldwide problem. Viruses can be written for, and spread on virtually any computing platform. When first introduced, they were often structured as boot sector attacks, typically promulgated by infecting floppy disks that are subsequently read during start-up. As the primary file transfer mechanisms today are through e-mail and other electronic means, boot sector viruses are no longer considered a major concern. Typically, viruses are written to affect client personal computers. However, if the personal computer is connected to other machines on a local area network, it is possible for the virus to invade these machines as well. Please refer to Section 6.6 (Malicious Code Protection) for detailed descriptions of the various types of malicious code, potential malicious code attacks and countermeasures, and requirements for malicious code detection.

### 7.2.2.1 Technology Overview

Malicious code scanning technologies prevent and/or remove most types of malicious code. The use of malicious code scanning technologies with current virus definitions is crucial in preventing and/or detecting attacks by all types of malicious code.

There are several basic categories of anti-virus technologies:

- Pre-Infection Prevention Products—a first level of defense against malicious code, used before a system has been attacked.
- Infection Prevention Products—used to stop replication processes and prevent malicious code from initially infecting the system.
- Short-Term Infection Detection Products—used to detect an infection very soon after the infection has occurred.
- Long Term Infection Detection Products—used to identify specific malicious code on a system that has already been infected for some time, usually removing the malicious code and returning the system to its prior functionality.

Please refer to Section 6.6.5.2 (Viruses and E-mail) for a more detailed description of the types of malicious code detection technologies.

### 7.2.2.2 General Considerations for Selecting the Technology

Malicious code protection should be applied for workstations with individual access to networks or information services, as well as for networks at the gateway (as discussed in

---

<sup>2</sup> The term “virus” is often misused as referring to *anything* that “infects” a computer *and* causes damage. A more appropriate term for any software that attacks a system is “malicious code.” Throughout the remainder of this section, the use of the general term *virus* is intended to encompass the broader class of malicious code and delivery mechanisms.



Section 6.4.2 (Malicious Code or Virus Detectors)). Malicious code can infiltrate and destroy data through network connections if allowed past the gateway or through individual user workstations. Although a single user can bring an infected disk to work, infecting his workstation and eventually the entire network, the majority of infections by malicious code result from file sharing. Today, many individual users keep all data files on networks or shared file systems instead of diskettes. Therefore continual protection of network connections at the gateway is a very important aspect of an overall security posture.

### 7.2.2.3 Important Features

When considering the selection of anti-virus technologies, there are a number of features that should be considered. This section identifies important features for selection. The section that follows discussed rationale for the selection of these features. Additional factors to consider when selecting a malicious code detection product can be found in Section 6.6.6 (Selection Criteria).

#### Detection Capabilities

- Data integrity checks.
- Does tool exploit malicious mobile code.
- Real time virus scanning.
- On-demand virus scanning.
- Different strains of polymorphic viruses.
- Viruses residing in encrypted messages, compressed files.
- Viruses in different languages (e.g., JAVA, ActiveX, Visual Basic).
- Trojan horses and worms.

#### Updates

- Can tool upgrade an existing version.
- Are regular updates available.
- Frequency of update releases.

#### Response Mechanisms

- Quarantine at the server level.
- Quarantine at the console level.
- Supply network-based responders.
- Send alerts to network or system administrators.
- Send alerts (in the case of e-mail borne viruses) to sender and receiver(s).

#### Platform Considerations

- What platforms the tool runs on.

- Does tool allow cross platform support.

## 7.2.2.4 Rationale for Selecting Features

When selecting anti-virus technologies, two important guidelines should be followed. The “best” technology may not be good enough by itself. Also, since data security technologies operate in different ways, one technology may be more useful than another in different situations. The following is rationale for evaluating features of specific technology offerings. Rating each of these categories will allow an organization to choose the best malicious code protection technology for their needs.

## Detection Capabilities

As discussed in Section 6.6.5.2 (Viruses and E-mail), most computer-virus scanners use pattern-matching algorithms that can scan for many different signatures at the same time. Malicious code detection technologies have to include scanning capabilities that detect known and unknown worms and Trojan horses. Most anti-virus technologies search hard disks for viruses, detect and remove any that are found, and include an auto-update feature that enables the program to download profiles of new viruses so that it will have the profiles necessary for scanning. The viral signatures these programs recognize are quite short—typically 16 to 30 bytes out of the several thousand that make up a complete virus. It is more efficient to recognize a small fragment than to verify the presence of an entire virus, and a single signature may be common to many different viruses.

Despite the fact that anti-virus applications are essential for the detection of known viruses, no mail filter or malicious code scanner can defend against a new mail worm attack. The recent “Love Bug” virus was caught quickly and still did a wealth of damage. And it seems to only be a matter of time before crackers figure out how to send email worms that infect systems without opening attachments.

## Updates

Maintaining an effective defense against virus and hostile code threats involves far more than the ability to produce perfect detection rates at a given point in time. With an average of nearly 300 new viruses discovered each month, the actual detection rate of anti-virus software can decline rapidly if not kept current. This anti-virus protection should be updated regularly. As new viruses are discovered, corresponding cures are developed to update protections. These updates should not be ignored. Anti-virus systems should do these updates automatically, reliably, and through a centrally controlled management framework. To stay current, these scanning programs must be updated when new viral strains are found and anti-viral codes written. This is why an enterprise-class anti-virus solution must be able to offer timely and efficient upgrades and updates across all client and server platforms.

## Response Mechanisms

Once malicious code has been detected, it must be removed. One technique is simply to erase the infected program, but this is a harsh method of elimination. Most anti-virus programs attempt to repair infected files rather than destroy them. If a virus-specific scanning program

detects an infected file, it can usually follow a detailed prescription, supplied by its programmers, for deleting viral code and reassembling a working copy of the original. There are also generic techniques that work well for known and unknown viruses. One method is to gather a mathematical fingerprint for each program on the system. If a program subsequently becomes infected, this method can reconstitute a copy of the original. Most tools perform scanning for viruses, but all do not detect and remove Trojan Horses, worms, and malicious mobile code upon all levels of entry. Most currently available anti-virus tools do not have the same capabilities when responding across a network. Additional countermeasures related to malicious code can be found in Section 6.6.4 (Potential Countermeasures).

The technology should be easy to use providing clear and uncluttered menu systems as well as meaningful screen messages. Help systems are essential, as users need current information regarding all types of malicious code. The trend is to provide online help; however, manuals should also be provided with the technology. The malicious code protection technology should be compatible with all hardware and software and should not create conflicts. The company that produces the technology should be stable and able to provide necessary local technical support for all questions and problems. The technology should be fully documented. That is, all messages and error codes should be deciphered and full installation guides and how-to manuals should also be provided.

## Platform Considerations

The computers to run this software must meet the hardware and software requirements specified by the manufacturer. The malicious code protection software should function properly and perform its duties without failing or interfering with other applications running on the same system.

### 7.2.2.5 Considerations for Deployment

Defense in depth dictates that any virus protection must be implemented across the enterprise. This means installing and managing anti-virus software on every system. Some advocate only installing anti-virus only on edge devices such as servers, firewalls, and gateways. Defense against viruses is only as good as its weakest link and if one system can be compromised, then the entire enterprise is at risk.

Centralized management for the anti-virus capabilities with a common set of policies is strongly recommended. Though some vendor offerings that cater to end-users being held responsible for security mandates, this can lead to more and varied security holes. What most often happens is that end users have their session interrupted with a pop-up screen that tells them their files are about to be scanned or that they are about to receive an anti-virus update. Many users tend to override the update manually, as it is distracting.

### 7.2.2.6 Considerations for Operation

Most anti-virus technologies provide a means for sending responses or alerts at the server level, and some at the console level. It is always desirable to notify anyone that may have been infected that malicious code has been detected. This should include system and network administrators. If it is encountered in e-mail transactions, it is desirable to notify the sender

and recipient. If it is found on a file system that knows the file owner, he should be notified. In general, anyone that could be notified should be.

## 7.2.3 Host Scanners—Host Vulnerability Scanners

As discussed above, on-line host monitoring technologies provide a critical layer of defense within enclave boundary protection. In addition to the host monitoring technologies, there is another class of technologies referred to as host scanners that can also be deployed to improve overall security posture. The Framework makes the distinction between these scanners and network monitoring devices. Monitors typically operate in near real time and tend to measure the effectiveness of the host's protection services as they are subject to attempted exploitation. This is somewhat of an "after the fact" measure, not a preventative measure. Scanners, on the other hand, are preventative measures. Typically, they operate periodically (or on demand) and examine hosts for vulnerabilities that an adversary could exploit, measuring security effectiveness.

The local environment is the logical place for addressing these host scanner technologies. Scanning can be performed at two levels. A *remote* (or network) scanner is run over a network against the target node by probing it for vulnerabilities. Here the software is running on an administrative system and scanning the target across the network. These network scanners are addressed in Section 6.5 (Network Scanners Within Enclave Boundaries) of the Framework. A *local* (or host) scanner runs ON a node as a software program that resides on the node itself. These scanners are discussed in this section of the Framework.

Unlike the near real time host monitoring technologies addressed above, host scanners are typically executed in a periodic or on-demand fashion, providing perspectives on the posture of a local environment. Section 8.2 (Detect and Respond as a Supporting Element) of the Framework provides a perspective on an overall Detect and Response infrastructure, however, since these assessments typically focus at the local level, they tend not to interact with, or be particularly relevant to a broader system infrastructure.

### 7.2.3.1 Technology Overview

Host-based vulnerability scanner tools examine the security posture of a host system from within the system, as opposed to network-based scanning tools that scan from the viewpoint of the network. They examine the contents of files looking for configuration problems. It compares what it finds against a set of pre-defined policies or best practices and generates alerts when it detects possible deficiencies in the system's security posture. These technologies catch security problems that are not visible at the network level and could be exploited by users with malicious intent that already have access, through valid (or otherwise, such as stolen authentication information) means, to the system.

## Detection

Scanners probe for weaknesses by comparing data about the host's configurations against a database of known vulnerabilities. Scanners work by either examining attributes of objects (e.g., owners and permissions for files) or by emulating an attacker. In this latter approach, the scanner runs a variety of scripts that try to exploit vulnerabilities in the host. Most scanners can

be configured to select which vulnerabilities to scan for and when the scans are to be performed. Some scanners allow an operator to incorporate their own scanning routines to look for site-specific application weaknesses. Some also offer capabilities for grouping hosts into scan groups, and for customized scan options by group.

## Scan Configuration Mechanisms

Each host in an enclave should be equipped with a host-based scanner. If the number of nodes is small, configuring the scanner and reviewing the results locally may be preferred, as network traffic overhead is minimized. If the network is large, it is often desirable to be able to configure one or more consoles to control distributed node scanners. Some technologies include software distribution frameworks for propagating this control. Hosts can be collected into groups, and one host can be a member of more than one group. Groups can be scanned at different times, and variations in the vulnerabilities inspected may be specified for each group. This enables the operator to scan some hosts “deeper” than others by tailoring the types of scans that are performed. For example, you can configure the scanners to search for user configuration errors on hosts that serve many users, and omit those scans on hosts (e.g., servers) that have no users.

## Response

When a host is scanned, some technologies develop a “fix script” with recommended corrective actions. It may be possible to customize this script, or to run it to eliminate the vulnerabilities identified by the scan. Some also provide an unfix script that enables the operator to undo the fix script.

### 7.2.3.2 General Considerations for Selecting the Technology

One advantage of periodic scanning is that resource utilization is less on the average than that required for real-time monitoring, since processing resources are required only when the scanner is active rather than on a constant basis. Unlike host monitoring technologies that are intended to catch adversaries in the act, scanners reveal weaknesses or vulnerabilities that could be exploited at some future time, whether they arise from accidental configuration errors or as a consequence of an intrusion or misuse. Since host scanners actually run on the target node, they can focus on problems that cannot be detected by remote (or network) scans. They can also inspect patch levels to ensure the latest security fixes have been installed. In the same vein, because scanners are run periodically, they are not able to detect events as they occur.

### 7.2.3.3 Important Features

When considering the selection of a host-based vulnerability scanners, there are a number of features that should be considered. This section identifies important features for selection. The section that follows discusses the rationale for selection of these features.

## Scanning Capabilities

- Does the tool offer the ability to add custom scanning routines to look for site- or technology-specific weaknesses of concern.

## Signature/Vulnerability Database

- Comprehensive list of vulnerabilities for the target host system.
- Periodic updates from the vendor.
- Entries can be added easily by user.
- Is the database backed by a vendor-funded research center, or just culled from Internet-based sources of vulnerability information, or some combination.

## Response Mechanisms

- Automatic shut off of vulnerable ports of entry.

## User Interfaces

- Can reports be viewed in real time.

## Reporting Capabilities

- Automatic alerting when new non-network ports are detected.
- Are all system answers logged in a database or file.
- Is there an updated database of network numbers with which to compare newly identified numbers.
- Does the database automatically combine logged information and place it in a report format.
- Does the tool provide suggested mitigation approaches for discovered vulnerabilities.

## Platform Compatibility

- What are the platforms (operating systems) upon which the tool will run.
- Does it use executables.
- Does it support scripts or macros.

## Source

- Has the tool been developed by the Government (or under Government sponsorship); if so it is reserved; can your organization obtain authorization for its use.
- Is the tool available from a reputable vendor.
- Is the tool in the public domain (e.g., freeware from the Internet); if so, is source code available.

## 7.2.3.4 Rationale for Selecting Features

### Scanning Capabilities

The type and level of detail of information provided among tools varies greatly. Although some can identify only a minimal set of vulnerabilities, others can perform a greater degree of analysis and provide detailed recommended mitigation approaches. Scanner technologies should be selected that cover the gamut of vulnerabilities for the given operating system environments (e.g., UNIX and Windows), including password vulnerabilities, access control, resource and file permission signatures, registry problems, and the like. Technologies that offer a comprehensive library of vulnerabilities periodically updated by the vendor should be selected. Capabilities including grouping of nodes into scan groups and customized scan options may be valuable for larger environments.

Some scanner technologies offer features that are useful depending on the training and skill levels of operators that will be using them. Depending on the planned usage of the scanner and the skills of the operators available, it is often desirable to select technologies that can be tuned to ignore some false positives. It is also desirable to select features that enable the scanner to be tuned for important application environments, such as database environments, Web Server environments, file server environments, firewalls, etc., since such profiles may differ for the function the system must provide to the enterprise.

### Signature/Vulnerability Database

A significant characteristic of host-based vulnerabilities is that they are extremely unique on a per-operating system, and even on a per application basis. In some cases, applications that are portable also port their vulnerabilities across platforms, and can have different vulnerabilities under different platforms. And obviously, operating structures differ drastically between the general Unix base (and their variants), the Windows 95/98 base, and the Windows NT/2000 base. It is important therefore, that the vulnerability database provided under the host-based IDS be comprehensive, adaptable, and well maintained by the vendor. The Framework strongly recommends selecting technologies in this area from vendors that maintain their own research capabilities, with specific expertise in operating system vulnerabilities, as opposed to those that simply incorporate vulnerabilities signatures culled from other Internet-based resources.

### Response Mechanisms

Assessment tools will continue to evolve in usability, with some vendors offering click-and-fix solutions. The assessment software flags vulnerabilities in terms of the risk posed to the network and the ease of the fix. Some technologies can generate trouble tickets to trigger a manual response. They may offer the ability to change policies in firewalls and other enclave boundary defense mechanisms. Some identify patches that should be installed. Some offer to obtain and install patches. Although installing patches is feasible, allowing the security administrator the ability to undertake these tasks and the difficulty of undoing configuration changes make this feature less desirable. Such features should be considered in light of an environment's existing configuration management policies and procedures.

## User Interfaces

Most scanners enable the operator to configure which vulnerabilities are to be scanned and when the scans are to occur. Typically, scanners are pre-configured with lists of vulnerabilities and can operate without customization. Some technologies allow operators to customize the vulnerabilities the scanner will investigate. More recently developed tools provide user-friendly front ends and sophisticated reporting capabilities.

## Reporting Capabilities

Usually the results are sorted into a file that can be accessed upon demand to review the results. Old technologies inundated customers with phonebook-sized reports on all the various vulnerabilities that the network faced. New technologies have database interfaces that prioritize vulnerabilities and allow network managers to deal with the network's problems in a logical manner. Many generate reports that are web-enabled with hot-links and other "labor savers." For sites with a few platforms running the scans and reading the reports on each node may be appropriate. For sites with large numbers of hosts, consideration should be given to consolidated reports available from a central server. If this feature is selected, the Framework recommends consideration for technologies that offer encryption for information transferred from the local hosts to the centralized server to protect the information about the vulnerabilities identified.

## Platform Compatibility

The computers to run this software must meet the hardware and software requirements specified by the manufacturer. The vulnerability scanner software should function properly and perform its duties without failing. Also, technologies should be considered to minimize the impact they may have on the performance of the host.

## Source

Host vulnerability scanner technologies are available from a variety of sources. Various organizations in the Government have developed their own tools, usually for use by members of specific communities. The quality of these tools typically varies based on the skills (and the testing) performed by the developing organization. If one of these is identified, authorization for its use will likely be required. Otherwise, tools are available commercially, and others can be downloaded over the Internet. The Framework recommends that if Government tools that have been tested and found to be effective, the selection of these technologies should focus on commercial tools available from reputable vendors. If a tool is downloaded from the Internet, it is important to determine if the source code is available, allowing them to be tested and evaluated by an experienced analyst. We strongly recommend against the use of tools downloaded from the Internet where no information about the source code is available. Since no reliable information about these tools is typically available, it is possible that the tool may not detect actual vulnerabilities, and worse yet, could actually introduce vulnerabilities (e.g., as a source of a malicious code attack).



### 7.2.3.5 Considerations for Deployment

It is often useful to consider deploying vulnerability scanners in conjunction with a host-based IDS. An IDS will be capable of identifying when a file is modified; however, it cannot determine what changes were made to that file. It can invoke a scanner as a response to inspect the contents of that file. For example, maintaining configurations of owners, groups, and permissions for files and directories is one typically challenging task. Scanners can ensure that these aspects of a security policy are properly implemented.

### 7.2.3.6 Considerations for Operation

It is important to consider what and when scans are performed. Otherwise, it is possible that mission critical servers become busy responding to simulated attacks during times of peak demand. Assessment frequency is a factor of how often network changes are made as well as the security policy for the enterprise. Depending on the organization, assessments may take place on a quarterly, monthly, weekly, or even daily basis. Some service providers offer scanning services on a subscription basis, ensuring that assessments take place regularly. It is recommended that features that provide automated vulnerability repair be disabled. If they are used, we strongly recommend the system is backed up fully (including all system and application software) prior to any automated repair.

## 7.2.4 Host Scanners—File Integrity Checkers

This is a specialized type of host scanner technology that verifies the integrity of files detecting when files have been changed. As with the host vulnerability scanner technologies discussed above, these technologies tend to run off-line, and thus are not a protection mechanism. Typically they operate periodically, based on an event (e.g., file access) or on demand.

### 7.2.4.1 Technology Overview

This is a small, tailored class of technologies that are configured with the location of specific key configuration files or executables (dependent upon the operating system in question) that are typically targeted by attackers attempting to compromise the system's integrity. These can include the registry environment, file permissions, security policy, account information, etc. The software typically generates cryptographic checksums of the targets and periodically checks to see if the files have been surreptitiously modified. The most widely known of these is Tripwire, but there have been some recent entries into the field, as well. A few host-host-based IDS monitors and host vulnerability scanners have integrated limited file integrity checking capabilities, and a number of technologies that have started out as integrity checkers are evolving into policy violations checkers and vulnerability scanners, so the two product lines are coalescing.

Most integrity checkers work using the same general paradigm. They operate on files identified from a library of known files to monitor. Dependent upon the platform and operating system in use, the technology creates unique identifiers typically based upon cryptographic checksums. These identifiers are then stored for future use. The file integrity program is subsequently executed, either automatically or manually, and new unique

identifiers are calculated. The integrity checker compares the new identifiers with the saved versions, and when a mismatch occurs, it notifies the operator or administrator that the file has been modified or deleted. The operator or administrator determines whether the differences indicate intrusive activity.

## 7.2.4.2 General Considerations for Selecting the Technology

In most cases, general considerations for use of file integrity checkers parallel those of host IDS and vulnerability scanning in general. A few additional discriminators do exist, however. Most important is that file integrity checkers are supported by cryptography, providing stronger protection against being defeated by intruders. If they are configured to run in near real time, file integrity checkers provide instantaneous indication of attack or failure. Also, if they are configured to run on files or data structures that do not change, their alarms require little or no interpretation. Unfortunately, file checkers also suffer from the same drawbacks as other host-based technologies, centered primarily on performance and resource consumption. It is also critical to ensure the baseline signatures from which the checkers function are both well protected from modification, and, in the case of dynamic configuration data structures, are created before the system is accessible to its user population. Table 7.2-2 below provides a synopsis of the advantages and disadvantages of File Integrity Checkers.

**Table 7.2-2 File Integrity Checker Considerations**

Advantages	Disadvantages
<p>Provides additional security protections based upon cryptographic methods</p> <p>Provides clear immediate evidence of intrusion when files that should <i>never</i> be modified are discovered modified vs. interpretation of host-based IDS reports and alarms</p> <p>System can operate within an encrypted environment since the host has access to decrypted versions of files.</p> <p>Systems can distribute the load associated with monitoring across available hosts on large networks.</p>	<p>Network activity is not visible to host-based sensors.</p> <p>Incurs additional resource overhead on the system, depending upon frequency of execution.</p> <p>Operating system vulnerabilities can undermine the integrity of host-based sensors and analyzers.</p> <p>File identifiers or signatures, even if based upon cryptographic checksums, must have their own, strong protection.</p> <p>Management and deployment costs associated with host-based systems are often greater than in other approaches to implementing an intrusion detection system.</p> <p>Host-based sensors are often platform-specific, which adds cost and required operator expertise.</p> <p>If not deployed before system is operational, may miss early system compromises.</p>

## 7.2.4.3 Important Features

When considering the selection of a host-based file integrity checking scanners, there are a number of features that should be considered. This section identifies important features for selection. The section that follows discusses rationale for the selection of these features.

## Scanning Capabilities

- Comprehensive set of files/data structures to monitor for each operating system (including data structure and directories environments such as Lightweight Directory Access Protocol (LDAP) or full X.500 services).
- Strong cryptographic checksums implemented as part of the identifier scheme.
- Centralized reporting for large enterprises.
- Built-in analysis or recommended action upon noticed modification.
- Self-checking.
- User can easily specify additional files/structures to monitor.

## Response Mechanisms

- Automated restoration of “clean” file or data structures.

## User Interfaces

- Does the tool have a graphical user interface for number entry, dialing status, and call results.
- Can reports be viewed in real time.

## Reporting Capabilities

- Automatic alerting when new, non-network ports are detected.
- Are all system answers logged in a database or file.
- Is there an updated database of network numbers with which to compare newly identified numbers.
- Does the database automatically combine logged information and place it in a report format.
- Does the tool provide suggested mitigation approaches for discovered vulnerabilities.

## Platform Compatibility

- What are the platforms (operating systems) upon which the tool will run.
- Does it use executables.
- Does it support scripts or macros.

### 7.2.4.4 Rationale for Selecting Features

## Scanning Capabilities

We strongly recommend technologies that offer a comprehensive library of files and data structures for tracking that is periodically updated by the vendor. As new vulnerabilities are discovered which include files or structures that an attacker might modify, the vendors

involved should provide immediate updates. Strong cryptography should be implemented as part of the checksum creation and re-check. Most scripted attack programs can already compensate for more widely known simple checksum hashing techniques and re-calculate checksums. Additionally, integrity checking technologies now exist for monitoring static portions of directory structures such as those found in Lightweight Directory Access Protocol (LDAP) or full X.500 Directory environments.

As with host vulnerabilities, file and data structures integral to any particular operating system environment are extremely unique on a per-operating system, and even on a per-application basis. In some cases, applications that are portable also port their vulnerabilities across platforms, and can have different vulnerabilities (characterized by different targeted files or data structures) under different platforms. And obviously, operating structures differ drastically between the general Unix base (and their variants), the Windows 95/98 base, and the Windows NT/2000 base. It is critically important therefore, that the database of files and data structures to monitor provided under the host-based integrity checker be comprehensive, adaptable, and well maintained by the vendor. We strongly recommend selecting technologies in this area from vendors that maintain their own research capabilities, with specific expertise in operating system vulnerabilities, as opposed to those that simply incorporate vulnerabilities signatures culled from other Internet-based resources.

## Response Mechanisms

Assessment tools will continue to evolve in usability, with some vendors offering click-and-fix solutions. This will be true in the file integrity-checking environment, as well, with some tools able to restore, from some secured back-up environment, files or environments that have been illegally modified.

## User Interfaces

Most file checkers enable the operator to configure which files and data structures are monitored and when the checks are to occur. Typically, they are pre-configured with lists of files and data structures to watch and can operate without customization. More recently developed tools provide user-friendly front ends and sophisticated reporting capabilities.

## Reporting Capabilities

Usually the results are sorted into a file that can be accessed upon demand to review the results. Old technologies inundated customers with phonebook-sized reports on all the various vulnerabilities that the network faced. New technologies have database interfaces that prioritize vulnerabilities and allow network managers to deal with the network's problems in a logical manner. Many generate reports that are web-enabled with hot-links and other "labor savers." For sites with a few platforms, running these checks and reading the reports on each node may be appropriate. For sites with large numbers of hosts, consideration should be given to consolidated reports available from a central server. If this feature is selected, the Framework recommends consideration for technologies that offer encryption for information transferred from the local hosts to the centralized server to protect the information about the possible intrusions or misuse identified by the unauthorized modifications.

## Platform Compatibility

The computers to run this software must meet the hardware and software requirements specified by the manufacturer. The vulnerability scanner software should function properly and perform its duties without failing. Also, technologies should be considered to minimize the impact they may have on the performance of the host.

### 7.2.4.5 Considerations for Deployment

Considerations for general use include understanding the periodicity with which to run the integrity re-check step, whether it should be done automatically or by operator command, and where the reports should be centralized. These are all dependent upon the sensitivity of the information being processed within the particular system or the criticality of the particular systems operations to the rest of the enterprise.

### 7.2.4.6 Considerations for Operation

The most important aspect to consider for integrity checker operation is deployment timing. To be their most effective, integrity checkers should be initialized on systems before they are placed into production and made generally accessible to their user communities. If they baseline monitored files and data structures anytime after a system has “gone live,” it is possible that the system has already become compromised and the integrity checker will miss changes that have already occurred. This is particularly the case in structures that are not supposed to remain static (e.g., access control databases vs. static executables that should not change from their installed release).

## 7.2.5 Discussion of Typical Bundling of Capabilities Within Products

At one point host monitors were offered as stand-alone devices. There are a number of offerings that combine these monitors with firewalls, routers, vulnerability scanners and the like as a means for vendors to leverage existing market positions to gain market share in related areas. Another trend that is becoming popular is for larger vendors to offer integrated architecture approaches, where they combine a number of related technologies as a bundled offering. They tend to prefer custom rather than standard interfaces to preclude the merging of other vendor offerings. This offers so-called “complete solutions,” however, it tends to lock the buyer into one particular product suite. While this may sound attractive, it is often valuable to be able to integrate various technologies together in order to take advantage of the detection capabilities available from the different implementations.

There is a natural linkage of these monitoring technologies with enterprise security management (ESM) systems. For several years, technology projections have indicated that host-based vulnerability assessment software will become a prime candidate for integration into system management platforms. Aspects of network-based products are also likely to find homes in network management platforms. While these views are prevalent, there remains little evidence that this has been, or will be the case for the immediate future.

## 7.2.6 Beyond Technology Solutions

While the focus of the IATF is on technology solutions, there are important operational aspects of effective network monitoring that are critical to an effective IA solution.

### Operational Planning

The Framework recommends the following guidance.

- Develop intrusion detection and antiviral-related requirements as an integral part of the enterprise security policy.
- Assess the ability of system administration personnel to perform intrusion detection and related vulnerability scanning.
- Consult with experienced intrusion detection and vulnerability scanning personnel regarding the best approach.
- Seek a balanced and symbiotic deployment of sensors.
- Consult with the appropriate legal council regarding affected personnel rights and procedures, as discussed below.
- Provide for adequate technical and legal training of all involved personnel.
- Acquire software and expertise from a high integrity vendor.
- Perform network monitoring consistent with the enterprise security policy.
- Tightly couple vulnerability scanning and intrusion detection activities.
- Intrusion Detection Activities
  - Look for intrusion evidence based on found vulnerabilities; use intrusion evidence to find and correct vulnerabilities.
  - Provide and monitor bogus sites/services/information. Possibly monitor intrusions through known vulnerabilities to satisfy prosecution requirements in conjunction with the appropriate legal authorities.
  - Perform intrusion responses that are approved by the appropriate authority.
- Network Malicious Code Detection Activities
  - Select and deploy virus scanning capabilities that are consistent with the location, functions and capabilities.
  - Acquire or download the appropriate anti-viral software from a high integrity source, and acquire any necessary hardware (such as an ancillary firewall dedicated to virus scanning incoming or outgoing traffic).
  - Institute enterprise wide anti-viral training and procedures.
  - Scan consistently based on time and/or events.
  - Follow up on all indications of potential contamination (as defined in the security policy and antiviral procedures for the enterprise).
  - Update antiviral software and hardware as appropriate, e.g., consistent with new releases of anti-viral software and specific experiences throughout the enterprise.

## General Activities

- Archive (within any legal constraints) audit and intrusion information and correlate with vulnerability scan information.
- Keep authorities apprised of all activities, ensuring that any legal rights are not violated.
- Continuously repeat steps, as appropriate.

## Privacy Concerns

Organizations may own the intellectual property of employees and may also legally restrict computer activities to only those approved by management. A common practice is to present this warning to all computer users as part of the normal login message. This does not mean that ALL managers in an enterprise own ALL of the transactions of ALL of the employees. Especially unclear is how to handle the conflict that arises between privacy and monitoring. Use of IDSs and system monitoring tools requires caution. Sniffers that search for key words in messages (such as “attack,” “weakness,” or “confidentiality”) as a standard set of *watchwords* may find them used in an appropriate manner depending on the type of correspondence. Audit trail reports may contain full command strings (including parameters). Knowing that an employee is sending several messages to a particular department (e.g., Human Resources) may be an infringement to his privacy. It is important to refer privacy concerns to the appropriate legal and policy organizations for the enterprise prior to deployment and use of these technologies.

### 7.2.7 For More Information

The list of reference materials used in the preparation of this section provide an excellent base of knowledge from which to draw on relevant technologies, there are a number of additional sources of information. This section of the Framework focuses on on-line sources since they tend to offer up-to-date information. These include the following.

#### 7.2.7.1 IA Technology Framework Executive Summaries

An important segment of the IATF is a series of “Executive Summaries” that are intended to provide summary implementation guidance for specific case situations. These offer important perspectives on the application of specific technologies to realistic operational environments. While these are still being formulated, they will be posted on the IATF web site <http://www.iatf.net> as they become available. [1]

#### 7.2.7.2 Protection Profiles

The National Security Telecommunications and Information Systems Security Policy (NSTISSP) No. 11 provides the national policy that governs the acquisition of IA and IA-enabled information technology products for national security telecommunications and information systems. This policy mandates that effective January 2001, preference be given to products that are in compliance with one of the following.

- International Common Criteria for Information Security Technology Evaluation Mutual Recognition Arrangement.
- National Security Agency (NSA)/National Institute of Standards and Technology (NIST) National Information Assurance Partnership (NIAP).
- NIST Federal Information Processing Standard (FIPS) validation program.

After January 2002, this requirement is mandated. Department of Defense (DoD) Chief Information Officer (CIO) Guidance and Policy Memorandum No. 6-8510, Guidance and Policy for Department of Defense Global Information Grid Information Assurance references this same NSTISSP No. 11 as an acquisition policy for the Department.

The International Common Criteria and NIAP initiatives base product evaluations against Common Criteria Protection Profiles. NSA and NIST are working to develop a comprehensive set of protection profiles for use by these initiatives. An overview of these initiatives, copies of the Protection Profiles, and status of various products that have been evaluated are available at the NIST web site <http://niap.nist.gov/>. [2]

### 7.2.7.3 Independent Third Part Reviewers of Relevant Vendor Technologies

ICSA Net Security Page [www.icsa.net](http://www.icsa.net)

Talisker's Intrusion Detection Systems [www.networkintrusion.co.uk/](http://www.networkintrusion.co.uk/)

Network Computing – The Technology Solution Center [www.nwc.com/1023/1023f12.html](http://www.nwc.com/1023/1023f12.html).

Paper on CMDS Enterprise 4.02 <http://www.Intrusion.com/Products/enterprise.shtml> (ODS Networks has changed its name to Intrusion.com).

Paper on CMDS Enterprise 4.02 <http://www.Intrusion.com/Products/enterprise.shtml> (ODS Networks has changed its name to Intrusion.com).

PC Week On-Line [www.zdnet.com/pcweek/reviews/0810/10sec.html](http://www.zdnet.com/pcweek/reviews/0810/10sec.html).

### 7.2.7.4 Overview of Relevant Research Activities

Coast Homepage – Perdue University [www.cs.purdue.edu/coast](http://www.cs.purdue.edu/coast).

UC Davis—<http://seclab.cs.ucdavis.edu>.

UC Davis—<http://www.gidos.org/>

UC Davis—<http://seclab.cs.ucdavis.edu/>

### 7.2.7.5 Overview of Selected Host Monitor Vendor Technologies

Axent Technologies [www.axent.com](http://www.axent.com)

cai.net <http://www.cai.net/>



## UNCLASSIFIED

Host-Based Detect and Respond Capabilities Within Computing Environments  
IATF Release3.0—September 2000

Cisco Connection Online [www.cisco.com](http://www.cisco.com)

CyberSafe Corporation [www.cybersafe.com](http://www.cybersafe.com)

Internet Security Systems [www.iss.net](http://www.iss.net)

Network ICE [www.networkice.com](http://www.networkice.com)

## Host Scanner Vendor Technologies

Axent Technologies [www.axent.com](http://www.axent.com)

cai.net <http://www.cai.net/>

Cisco Connection Online [www.cisco.com](http://www.cisco.com)

CyberSafe Corporation [www.cybersafe.com](http://www.cybersafe.com)

Internet Security Systems [www.iss.net](http://www.iss.net)

Network ICE [www.networkice.com](http://www.networkice.com)

# References

1. Information Assurance Technical Framework (IATF) <http://www.iatf.net>.
2. National Institute of Standards and Technology <http://niap.nist.gov/>.

# Additional References

- a. Amoroso, Edward, Intrusion Detection. Intrusion.Net Books. 1999.
- b. Escamilla, Terry. Intrusion Detection, Network Security Beyond the Firewall. Wiley Computer publishing. 1998.
- c. Northcutt, Stephen. Network Intrusion Detection, An Analyst's Handbook. New Riders Publishing. 1999.
- d. Snapp, Steven R., et al. A System for Distributed intrusion Detection. IEEE CH2961-1/91/0000/0170. 1999.
- e. Balasubramaniyan, J. S., et al. An architecture for Intrusion Detection using Autonomous Agents. COAST Technical Report. 11 June 1998.
- f. AXENT Technologies, Inc. Intruder Alert 3.5 IDS Review Guide, May 2000.
- g. AXENT Technologies, Inc. Everything You Need to Know About Intrusion Detection, 1999.
- h. Schneider, Sondra, et al. Life After IDS. Information Security Magazine. Volume 2, Number 9. September 1999.
- i. Graham, Robert. New Security Trends for Open Networks. SC Magazine. October 1999.
- j. SC Magazine. Intrusion Detection. June 2000.
- k. Information Assurance Technology Analysis Center (IATAC). Tools Report on Intrusion Detection. Defense Technical Information Center. December 1999.
- l. Maes, V. How I Chose an IDS. Information Security Magazine. Volume 2, Number 9. September 1999.
- m. Concurrent Technologies Corporation. Attack Sensing, Warning, and Response (ASW&R) Trade Study Report Intrusion Detection System. Report No. 0017-UU-TE-000621. April 14, 2000.
- n. Information Assurance Technology Analysis Center (IATAC). Tools Report on Vulnerability Analysis Information. Defense Technical Information Center. March 15, 2000.
- o. Ulsch, Macdonnell and Joseph Judge. Bitter-Suite Security. Information Security Magazine. Volume 2, Number 1. January 1999.
- p. Concurrent Technologies Corporation. Attack Sensing, Warning, and Response (ASW&R) Baseline Tool Assessment Task Anti-Virus Trade Study Report. Report No. 0017-UU-TE-000623. April 13, 2000.

**UNCLASSIFIED**

Host-Based Detect and Respond Capabilities Within Computing Environments  
IATF Release3.0—September 2000

- q. Department of Defense (DoD) Chief Information Officer (CIO) Guidance and Policy Memorandum No. 6-8510, Guidance and Policy for Department of Defense Global Information Grid Information Assurance.
- r. National Security Telecommunications and Information Systems Security Policy (NSTISSP) No. 11. National Policy Governing the Acquisition of Information Assurance (IA) and IA-Enabled Information Technology (IT) Products. January 2000.

**UNCLASSIFIED**

Host-Based Detect and Respond Capabilities Within Computing Environments  
IATF Release 3.0—September 2000

**This page intentionally left blank.**

**UNCLASSIFIED**

7.2-33

## Chapter 8

# Supporting Infrastructure

A principal tenet of the Defense in Depth philosophy is to provide defenses against cyber intrusions and attacks, and deal effectively with and recover from attacks that penetrate those defenses. The supporting infrastructures are a set of interrelated activities and infrastructures providing security services to enable and manage the framework's technology solutions. Currently, the defense-in-depth strategy defines two supporting infrastructures:

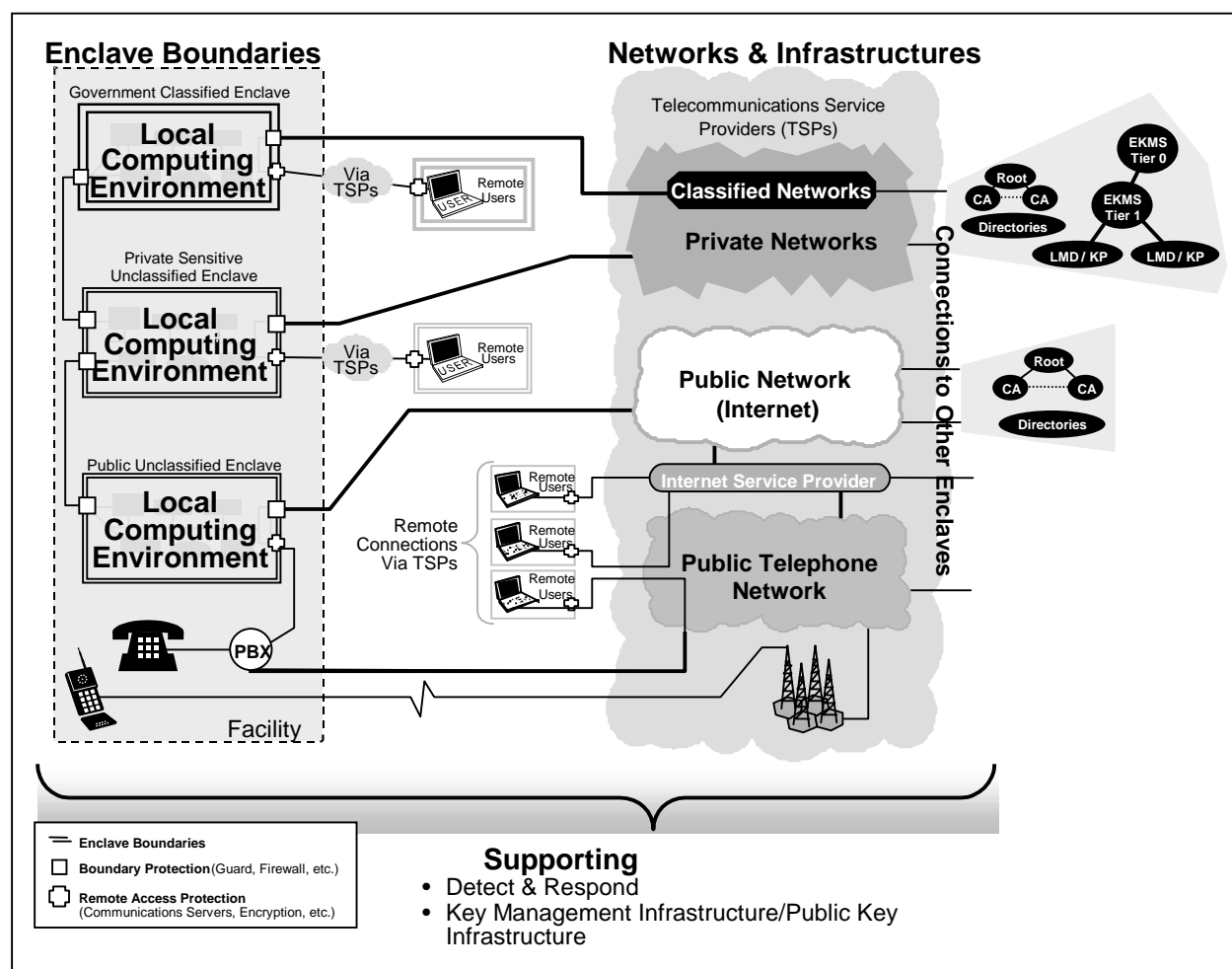
- *Key Management Infrastructure/Public Key Infrastructure (KMI/PKI)* for the generation, distribution, and management of security credentials like keys, and certificates; and
- *Detect and Respond* for providing warnings, detecting and characterizing suspected cyber attacks, coordinating effective responses, and performing investigative analyses of attacks.

Today's information infrastructures are not sufficiently secure to provide the full range of services needed to defend against the threats anticipated for the Global Information Grid (GIG). Thus, the defense in depth strategy provides overlays of IA features to realize an effective defense. Key Management (including Public Key Management) is fundamental to many IA protection technologies. Since our ability to provide airtight protection is neither technically nor economically feasible, we must re-enforce those protection technologies with capabilities to detect, respond to, and recover from cyber attacks that penetrate those protections.

Cryptography-enabled services rely on a key management or public key infrastructure to provide a trustworthy foundation. The KMI/PKI supporting infrastructure focuses on the technologies, services, and processes used to manage public key certificates and symmetric cryptography. As shown in Figure 8-1, the KMI/PKI infrastructure touches almost all portions of the networked environment.

KMI/PKI hardware and software at the enclave level provide local authorities (e.g., KMI managers) with capabilities to order and manage KMI/PKI products and services, issue certificates, and generate traditional symmetric keys. KMI at the Wide Area Network (WAN) level provides certificate, directory, and key generation and distribution functions.

The PKI strategy is based heavily on multiple levels of assurance since it is not cost effective to provide high-assurance protection for all PKI-enabled services. High assurance is needed when public key capabilities are used as the primary means to protect national security information. For other services, a medium assurance PKI is appropriate based on commercial technology. The medium assurance PKI will initially use software-based end-user tokens, but will evolve to the use of hardware tokens.

**Figure 8-1 Supporting Infrastructures: KMI/PKI**

Since one of the major features of a PKI is to provide widespread interoperability, and there is a broad base of non-interoperable commercial PKI technology solutions on the market today, we recommend a foundational PKI be fielded quickly, so that other efforts can build upon it. The PKI should support interoperability with external federal, foreign, and public domains. One way to achieve interoperability is through cross certification. Further study is required to decide where cross certification is best used. With PKI technology still immature and changing rapidly, the strategy for fielding a large-scale PKI quickly should be to make it a simple infrastructure that provides only basic cryptographic capabilities, including digital Identifications (IDs), compromise recovery, key recovery, archive and the like. Departments, agencies, and corporations are then free to build atop this infrastructure for such capabilities as access control.

It is not clear whether the higher assurance PKI is best operated by corporate personnel or outsourced. A number of Government organizations (including a major effort by the DoD) deploy and operate PKI pilots to gather operating information to evaluate its impact on mission (and business) performance and assess whether portions should be outsourced.

The local environments will maintain the option of deploying sensors, and possibly analysts to interpret the results of, and, when appropriate, react to the implications of these outputs. Beyond the local environment, each organization, or perhaps community has to determine what information should be reported, in what format, under what situations, and to whom.

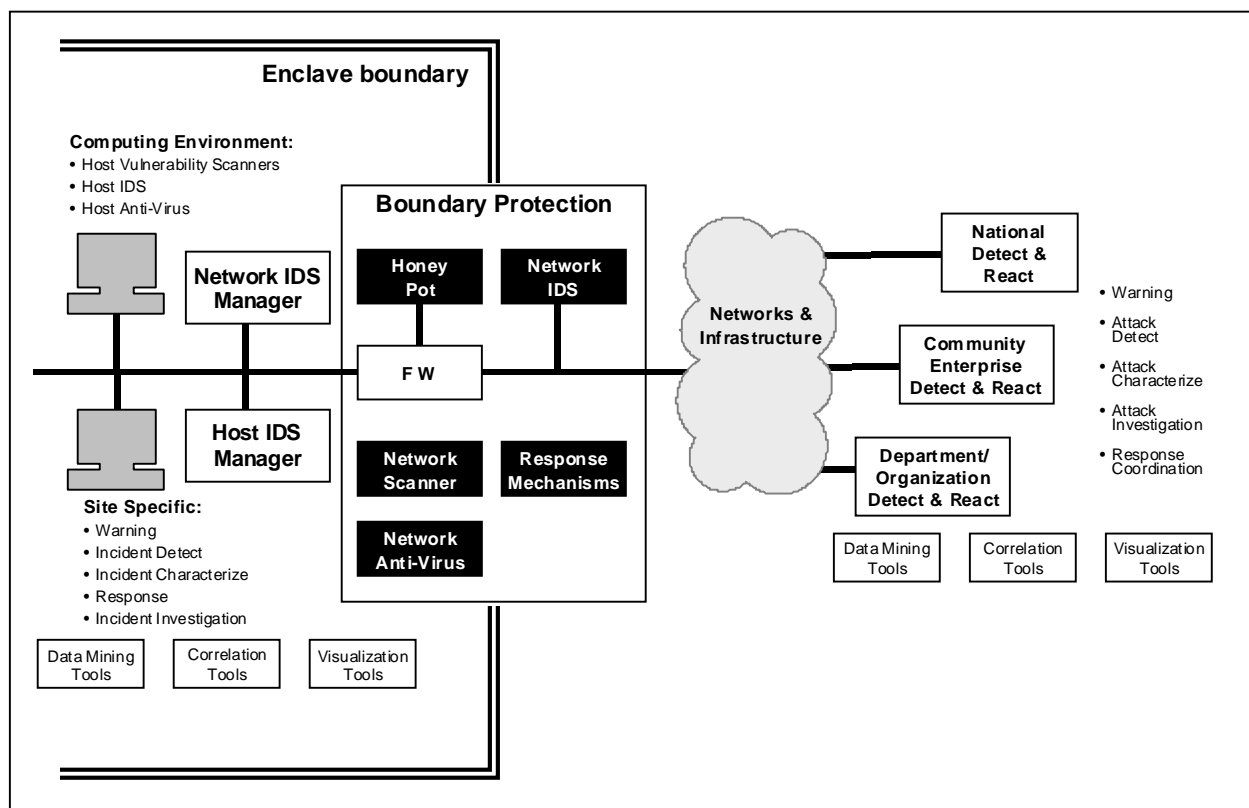
While planning for a Detect and Respond infrastructure, it is important to recognize that the enterprise networks and systems that it will support must also be structured to provide information to, and take advantage of the services and information such as infrastructure provides. This section provides good engineering practices for an enterprise to enhance its Detect and Respond capability.

When considering a general construct for a Detect and Respond infrastructure, a primary consideration is the perspective that the infrastructure will provide for its support. The reality is that most infrastructures are inherently hierarchical by their nature, and this one is no exception. What tends to be the case is that information about incidents, which is usually sensed at the lowest layer in the hierarchy, is promulgated up to higher layers with some form of reporting. Warning and response coordination that are more typically derived from higher layers are disseminated from those higher layers down.

There is a wide range of functions needed to support Detect and Respond, and technology solutions are not available to perform many of these functions automatically. Thus, many of these functions are performed by analysts, network operators, and system administrators who apply basic support technologies to ease their tasks. To deal with this from a technology viewpoint, we identify the functions that these analysts (and their tools) are attempting to perform, and then discuss the technologies that are available to realize these functions.

The Detect and Respond infrastructure element provides the functional and management capabilities to provide warning alerts of possible upcoming cyber attacks, and to assist local environments to detect, characterize, respond to and recover from attacks. Figure 8-2 highlights the areas of the high-level Defense Information Infrastructure (DII) context that comprise the detect and respond infrastructure.

Since the local environments are the logical location for sensors, the network-based sensor functions are discussed in Chapter 6 (Defend the Enclave Boundary/External Connections) and their host-based counterparts are covered in Chapter 7 (Defend the Computing Environment). We recognize that local environments have the option to implement as much or as little as they feel is prudent, obtaining services and support from the infrastructure. Detect and Respond processes and functions in the context of the supporting infrastructure are the focus of this section.



**Figure 8-2 Supporting Infrastructures: Detect and Respond**



## 8.1 Key Management Infrastructure/ Public Key Infrastructure (KMI/PKI)

This section focuses on management of the Supporting Infrastructure. Following introductory tutorial information, PKI Certificate management, Symmetrical Key management, Directory management, and Infrastructure management will be highlighted. Each of the process discussions is self-contained so that the reader can review only those KMI/PKI services and processes that are of interest. They will include specific requirements applicable to that process and KMI/PKI service, important threats and countermeasures, and the range of technologies used to implement the process. Table 8.1-3 defines at a high level, the way each process relates to the various KMI/PKI services. The remainder of the section will present a range of KMI/PKI solutions used by or planned for protected networks.

### 8.1.1 KMI/PKI Introduction

KMI/PKI is unique in the framework because it does not directly satisfy subscriber's security requirements, but forms building blocks used by other security technologies. The KMI/PKI is an enabler; however, the KMI/PKI architecture is heavily dependent on the specific applications it supports. Table 8.1-1 takes the subscriber categories described in Chapters 5 (Defend the Network and Infrastructure) and 6 (Defend the Enclave Boundary/External Connections) of the framework and relates them to the required KMI/PKI services. For example, a Virtual Private Network (VPN) provides an encrypted pipe between two enclaves. The KMI/PKI infrastructure supplies keys and certificates to the cryptographic devices that provide authentication and encryption. Additional services could include key recovery and a directory to provide access to subscriber's public certificates.

**Table 8.1-1 KMI/PKI Services Support to Subscriber Categories**

Subscriber Categories	KMI/PKI Services			
VPN	Key generation	Certificate management	Key recovery	Directory
Network Access	Key generation	Certificate management	Value-Added Services	Directory
Remote Access	Key generation	Certificate management	Key recovery	Directory
Multi-Level Security	Key generation	Certificate management	Directory	

A second area, where KMI/PKI differs from the Framework's other solutions, is that it distributes its security throughout a number of separate elements. These elements require extensive security, e.g., encryption, certificate management, compromise recovery, among themselves to protect the subscriber's key or certificate. Because of the repercussions of a successful attack against the KMI/PKI, internal infrastructure security requirements are often more stringent

than those required by subscriber applications. There are unique requirements on the infrastructure, e.g., policy management, and the level of security assurance for infrastructure components is usually higher than for subscriber applications.

### 8.1.1.1 KMI/PKI Services

Current KMI/PKI implementations generally consist of several stovepipe infrastructures from different organizations, supplying different subscriber solutions. The end subscriber may need support from several of the stovepipes for a single application. Today, subscribers have to contact each infrastructure separately to get service. High cost, dwindling manpower, and higher subscriber expectations are pressuring a merger of these stovepipes into larger infrastructure elements supporting multiple subscriber requirements.

This chapter will discuss four of the operational services supplied by the KMI/PKI supporting infrastructure. These KMI/PKI services support many subscriber applications and consequently employ different (but related) mechanisms and have unique security requirements. The first two services describe functions that directly support subscriber applications. The last two services are functions required by the subscriber functions to work properly.

The first KMI/PKI service is symmetric key generation and distribution. This is still the primary key management mechanism within the Government classified community. The banking community with their extensive use of the Data Encryption Standard (DES) encryption is another major user of symmetric key management. While symmetric key is being replaced by asymmetric key agreement in many applications, it has application outside the Government classified community in such areas as multicast and low-bandwidth applications such as wireless. Symmetric key management is a process where a central element (it could be one of the subscribers or a trusted independent element) generates, distributes, and manages a “secret key” for multiple recipients. Each recipient uses the same secret key for security processing between themselves and the other recipients for the life of the key.

The second KMI/PKI service is support for asymmetric cryptography (often called *public key cryptography*) and its associated certificate management. Asymmetric cryptography usually employs digital certificates to allow subscribers to authenticate the public portion of the asymmetric cryptography public/private key pairs. This authentication is important, because the security services that asymmetric cryptography provides depend on the subscriber of a public key (called the *relying party*) being assured that the public key is associated with a specific identified subscriber. Digital certificates (often called X.509 certificates, after the international standard which defines their format) cryptographically bind identities to public keys. Together, the components, personnel, facilities, services and policies that are used to generate and manage public key certificates define a *public key infrastructure*. Public key infrastructures can generate and manage digital signature certificates (used for authentication, data integrity, and non-repudiation) as well as key management certificates (used for confidentiality). The commercial community relies heavily on public key cryptography, and commercial vendors offer a wide variety of public key infrastructure products and services.

The third KMI/PKI service is directory service. Directory servers provide access to the public information required with PKI such as the public certificate, the related infrastructure certificates, and the compromised key information. Directory services can be provided either by a global set of distributed directories, e.g., X.500 Defense Message System (DMS) directories, or

by an on-line repository at a single site. Directories are normally very closely coupled with PKI, but are also used for other services.

The final KMI/PKI service is managing the infrastructure itself. The other infrastructure architectures discussed in this section consist of a number of elements working together to provide the subscriber service. The distributed nature of the infrastructure places additional functional and procedural requirements on the KMI/PKI and the sensitivity of the application places additional security requirements on the KMI/PKI. The internal structure of the infrastructure varies with the application(s) it supports. For example, the level of assurance demanded by the applications dictates many of the internal aspects of the KMI/PKI.

### 8.1.1.2 Security Applications

The security applications supported by the KMI/PKI differ depending on the type of cryptography that is being used by the application. Symmetric cryptography primarily provides confidentiality services both for data transmission and storage. It can also support other mechanisms such as Transmission Security (TRANSEC), e.g., spread spectrum, or in combination with additional mechanisms, data integrity and authentication during data transmission. Public key cryptography in conjunction with certificate management provides the full range of security services. Unlike symmetric cryptography, it can provide authentication and integrity for both data transmission and data storage. While it can encrypt information, this is an extremely inefficient process and is normally provided by a symmetric algorithm. Table 8.1-2 describes the security applications that each type of cryptographic algorithm supports.

**Table 8.1-2 Security Applications Supported By Cryptographic Type**

Security Applications	Symmetric Cryptography	Asymmetric Cryptography
Authentication	*	<b>X</b>
Non-repudiation	*	<b>X</b>
Transmission Confidentiality	<b>X</b>	
File Encryption	<b>X</b>	
Integrity	*	<b>X</b>
Availability (e.g., Spread Spectrum)	<b>X</b>	
Key Agreement		<b>X</b>
* These services can be enabled by symmetric cryptography when provided in conjunction with other mechanisms, e.g., a Cyclic Redundancy Check (CRC) encrypted with the message.		

### 8.1.1.3 Infrastructure Process

The KMI/PKI consists of a number of processes that all have to work together correctly in order for the subscriber service to be secure. Each of these processes is necessary at some level in all KMI/PKI architectures. These processes are listed below.

- *Registration* – Enrolling those individuals that are authorized to use the KMI/PKI.

## UNCLASSIFIED

Key Management Infrastructure/Public Key Infrastructure (KMI/PKI)  
IATF Release 3.0—September 2000

- *Ordering* - Request for the KMI/PKI to provide a subscriber either a key or a certificate.
- *Key Generation* - Generation of the symmetric or asymmetric key by an infrastructure element.
- *Certificate Generation* - Binding of the subscriber information and the asymmetric key into a certificate.
- *Distribution* - Providing the keys and certificates to the subscribers in a secure, authenticated manner.
- *Accounting* - Tracking the location and status of keys and certificates.
- *Compromise Recovery* - Removal of compromised keys and invalid certificates from the system in an authenticated manner.
- *Rekey* - The periodic replacement of keys and certificates in a secure, authenticated manner.
- *Destruction* - Destruction of the secret key when it is no longer valid.
- *Key Recovery* - Ability to a recover subscriber's private encryption key without direct access to the subscriber's copy of the key.
- *Policy Creation* - Definition of the requirements for employment of the previous processes.
- *Administration* - Running the infrastructure.
- *Value-added PKI Processes* - PKIs may also support optional value-added processes including archive, time-stamp, and notary services. Because all PKI architectures do not support these features, this section will not discuss them further.

The complete set of KMI/PKI processes are usually allocated to several elements performing independent tasks that require extensive coordination between elements. For most of the processes there are a number of ways of implementing the services based on the application supported; the security required; and the cost (money, people, performance) the subscriber is willing to pay. Each process contributes to the overall security of the KMI/PKI and has different forms of attacks and countermeasures. Table 8.1-3 defines the basic requirements for implementing each process for the four KMI/PKI services. Figure 8.1-1 depicts the interaction of these services.

**Table 8.1-3 KMI/PKI Processes**

<b>Processes</b>	<b>Certificate (public key) Management Section 8.1.2</b>	<b>Symmetric Key Management Section 8.1.3</b>	<b>Infrastructure Directory Services Section 8.1.4</b>	<b>Infrastructure Management Section 8.1.5</b>
Policy Creation	N/A	N/A	N/A	Define the domain's policy and method for enforcing the policy
Registration	Registration of people who can authorize subscribers	Registration of people authorized to order key	Registration for people authorized to update directory	Define process of authorizing changes to the infrastructure's trust model, e.g., new elements, cross-certification

# UNCLASSIFIED

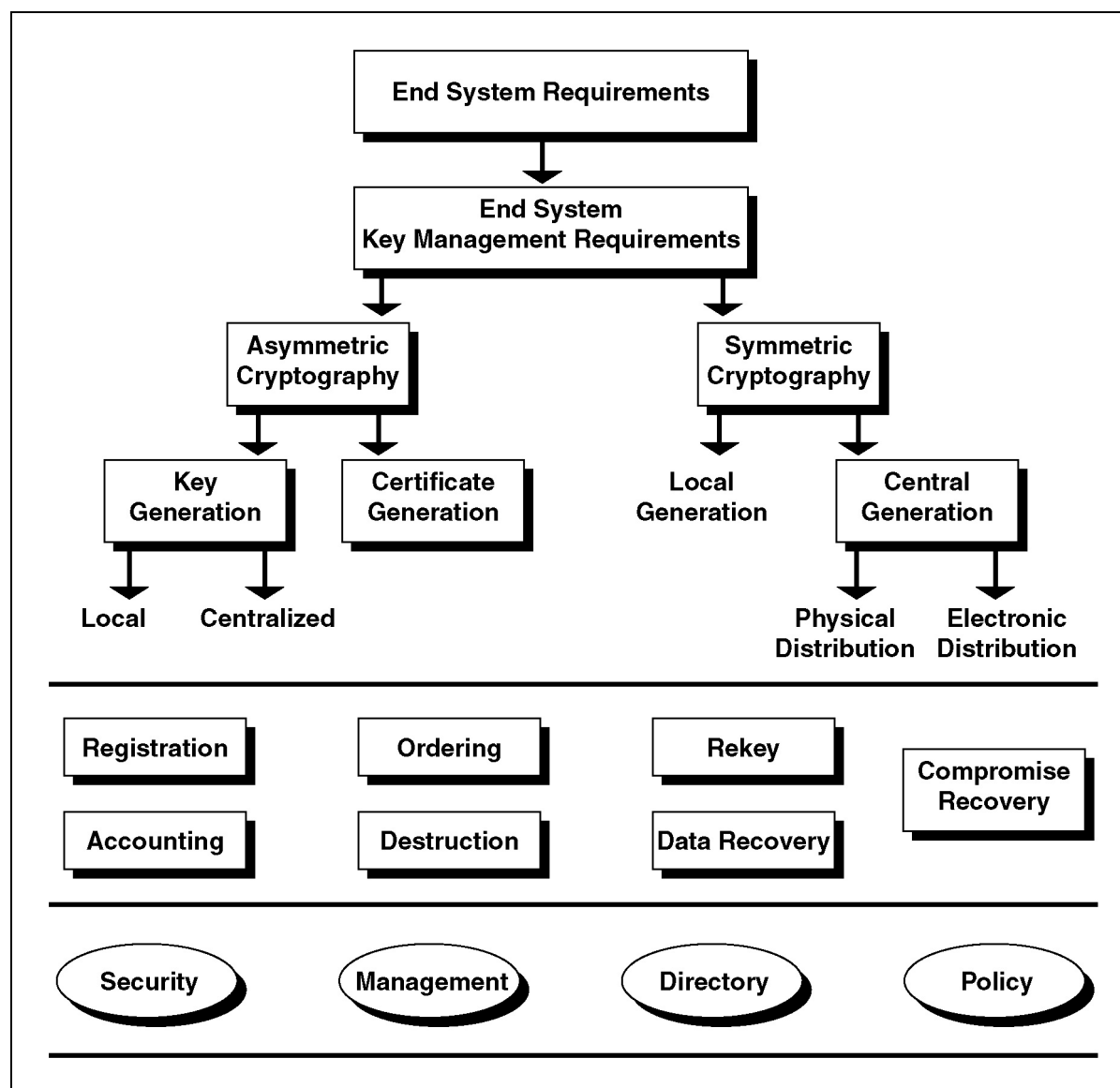
Key Management Infrastructure/Public Key Infrastructure (KMI/PKI)  
IATF Release 3.0—September 2000

<b>Processes</b>	<b>Certificate (public key) Management Section 8.1.2</b>	<b>Symmetric Key Management Section 8.1.3</b>	<b>Infrastructure Directory Services Section 8.1.4</b>	<b>Infrastructure Management Section 8.1.5</b>
Ordering and Validation	<ol style="list-style-type: none"> <li>1. Validation of the information in the certificate</li> <li>2. Validation of the key generation request</li> <li>3. Receiving the public key</li> </ol>	Validation of order	Validation of the information request	<ol style="list-style-type: none"> <li>1. Validation process for changes to the trust model</li> <li>2. Receiving the public key of the infrastructure elements</li> </ol>
Generation	<ol style="list-style-type: none"> <li>1. Generation of the public/private key pairs</li> <li>2. Generation of the certificate</li> </ol>	Generation of key	Adding information to the directory	<ol style="list-style-type: none"> <li>1. Generation of the root public/private keys</li> <li>2. Generation of the root certificate</li> <li>3. Generation of the infrastructure elements public/private keys</li> <li>4. Generation of the infrastructure elements certificates</li> <li>5. Generation of the cross-certificates</li> </ol>
Distribution	<ol style="list-style-type: none"> <li>1. Providing the certificate to the subscriber</li> <li>2. Validation that the person getting the certificate has the private key corresponding to the bound public key</li> <li>3. Providing the Policy Approving Authority (PAA) public certificate to the subscriber in an authenticated manner</li> </ol>	<ol style="list-style-type: none"> <li>1. Delivery of the key to the Custodian</li> <li>2. Loading the key into the cryptographic device</li> </ol>	Providing information to subscriber	<ol style="list-style-type: none"> <li>1. Provide the root certificate to each infrastructure element in an authenticated manner</li> <li>2. Provide each element with their certificates</li> <li>3. Validate that each infrastructure element has the private key corresponding to the public key</li> <li>4. Provide each element with the domain's cryptographic parameters in an authenticated manner</li> </ol>
Compromise Recovery	<ol style="list-style-type: none"> <li>1. Provide Compromise Key List (CKL) of compromised keys</li> <li>2. Provide on-line validation of the live-ness of certificates</li> </ol>	Supersession of all devices using the compromised key	Fixing a hacked directory	Procedures for reconstituting the infrastructure in case of disaster or compromise of any infrastructure element
Accounting	Tracking the location and status of key & certificates throughout life-cycle	Tracking the location and status of key throughout life-cycle	Auditing who makes changes to the information in the directory	Ensuring that the infrastructure elements operate within the policies and procedures defined by the PAA
Key Recovery	Appropriate key recovery mechanisms	N/A	N/A	Root signature key might need key recovery?
Rekey	<ol style="list-style-type: none"> <li>1. New certificate</li> <li>2. New key</li> </ol>	Rekey of the cryptographic device	N/A	Process for changing the root key(s)

# UNCLASSIFIED

Key Management Infrastructure/Public Key Infrastructure (KMI/PKI)  
IATF Release 3.0—September 2000

Processes	Certificate (public key) Management Section 8.1.2	Symmetric Key Management Section 8.1.3	Infrastructure Directory Services Section 8.1.4	Infrastructure Management Section 8.1.5
Destruction	Zeroize private key at the conclusion of use	Zeroize key at the conclusion of the cryptoperiod	Removing in- formation from the directory	Zeroize the infrastructure elements private key at the conclusion of use
Administration	N/A	N/A	N/A	Procedures for operating the infrastructure securely and enforcing the system policies



**Figure 8.1-1 Interactions of the KMI/PKI Applications Operational Services**

## 8.1.1.4 Requirements

This section includes both subscriber and infrastructure requirements. Because of the variety of issues involved in KMI/PKI, no single set of requirements can be consistent and complete for all applications. This paragraph outlines some of the high level requirements. It consists of both functional and operational requirements. Unlike most of the subscriber requirements identified in the Framework, the KMI/PKI has a large operational component. Once initialized, most subscriber solutions need little or no subscriber interaction, e.g., once the VPN has deployed the cryptographic device, the only update is the KMI/PKI task of rekeying periodically. The KMI/PKI on the other hand, generally requires extensive human interaction throughout its processing. This close coupling of people and service place additional requirements on the KMI/PKI that has implications on the security solution.

### 8.1.1.4.1 Subscriber Requirements

#### Symmetric Cryptography

- The key comes from an approved, authorized, authenticated source.
- The key is protected during distribution.

#### Asymmetric Cryptography

- The subscriber or the KMI/PKI shall generate the public/private key pair.
- The certificate information is accurate, current, and reflects a valid association with a uniquely identified subscriber.
- The certificate binds the public key associated with the subscriber's private key with the subscriber's identification.
- The trusted element's certificate is distributed to the subscriber in an authenticated manner.
- The subscriber can determine the current status of certificates in a timely manner.
- The KMI/PKI only provides a copy of a private key to authorize data recovery entities as defined by policy, e.g., subscriber or subscriber's organization.

### 8.1.1.4.2 Infrastructure Management Requirements

#### Symmetric Cryptography

- Ensures that requests for key generation or distribution only come from authorized sources.
- Key generation is secure and robust.
- The delivery mechanism protects the key from compromise.
- Key is only distributed to authorized subscribers.
- The system accounts for key during its entire life cycle (ordering, generation, distribution, use, rekey, and destruction).
- The infrastructure removes compromised keys from the system.

## Asymmetric Cryptography

- Ensures that a request for a certificate comes from an authorized source.
- Prior to generating the certificate, the system ensures that the information in the certificate corresponds to the requesting subscriber.
- The Certification Authority (CA) places the correct public key into the certificate.
- If the infrastructure generates the private key agreement key, it is generated securely and securely transmitted to the subscriber.
- The infrastructure must ensure integrity and provide its certificates in an authenticated non-repudiated manner to each subscriber.
- The infrastructure must provide compromise information to subscribers in a timely manner.
- The infrastructure must ensure high assurance in the registration of infrastructure elements.
- The system accounts for the life cycle of key (ordering, generation, distribution, application, rekey, destruction, and archive).
- The key recovery mechanism of the KMI/PKI only provides access to the private key to authorized entities, e.g., subscriber's organization.
- The key must be protected by the key recovery mechanism of the KMI/PKI during storage.
- The recovered key must be protected during distribution to the subscriber.

### 8.1.1.4.3 Interoperability Requirements

Note: Interoperability of the Key Management cryptographic infrastructure does not guarantee subscriber application interoperability.

## Symmetric Cryptography

- Keys and compromise information can be distributed to all subscribers.
- Format of the key must be the same for all subscribers.
- Algorithms and initial parameters must be the same for all subscribers.

## Asymmetric Cryptography

- When cross-certifying, the policies must be approved by each PKI.
- The subscriber may need to accept certificates from multiple domains.
- The infrastructure may need to support multiple algorithms and offer the subscriber the choice of algorithm to sign the certificate.
- The format of the keys and certificates must be the same for all subscribers, e.g., certificate profiles, use of X.509.
- Algorithms and initial parameters must be the same for all subscribers.
- Compromise recovery information must be available to all subscribers.



## 8.1.1.5 Attacks and Countermeasures

The goal of any attack against the infrastructure is to use it as a basis for attacking a subscriber's environment. Attacking the infrastructure does not provide an adversary with the subscriber's information (beyond audit information that may be archived), but may be used as a basis for a further attack against the subscriber. An attacker may directly target the information provided by the infrastructure, e.g., symmetric key, certificate, or may attack the infrastructure elements in order to later attack a subscriber, e.g., place a Trojan horse in an infrastructure element to substitute a known key for the subscriber's valid key. Table 8.1-4 lists several interesting attacks and potential countermeasures.

**Table 8.1-4 Attacks and Countermeasures**

<b>Attacks Against User via Infrastructure Support</b>	<b>Attacks Against Infrastructure</b>	<b>Countermeasures</b>
Read traffic due to weak cryptography (compromised, weak keys). Masquerade (get a certificate with false information). Denial of service (prevent signature from verifying, e.g., attack directories). Man-in-the-middle attack.	Violate trust model, e.g., generate an unauthorized cross-certification. Acquire unauthorized certificate, e.g., insider, incorrect identification. Force subscriber to have weak key, e.g., known key, failed randomizer. Denial of service - attack directories. Compromise key during distribution. Unauthorized access to key recovery key. Compromise Personal Identification Number (PIN) to gain access to subscribers private key (generation, distribution, use). Prevent subscriber from determining compromise status during validation. Substitute the attacker's public key for the subscriber's public key. Place malicious software into infrastructure elements. Cryptanalytic attack against the PKI's private keys.	Use security features of the protocols, e.g., name constraints, policy mapping. Proper management of the infrastructure. Multi-person control on the certificate approval and generation process. Protected distribution, e.g., benign fill. Robust compromise recovery. Use tokens to generate and protect private keys. High assurance operating systems in infrastructure components. Strong authentication on infrastructure services, e.g., directories and key recovery. Coordinate certificate request content with the Security Officer, Personnel Officer, Authorization Officer, and Privilege Assignment Officer. Independently certify the content of certificates against the officially approved certificate requests.

## 8.1.2 Certificate Management

One of the primary functions of KMI/PKIs is the generation, management, and distribution of asymmetric key material and certificates used within a variety of public key-based applications. The portion of the KMI/PKI dedicated to the management of keys and certificates is the PKI. This section provides an overview of the architecture and the processes or functions associated

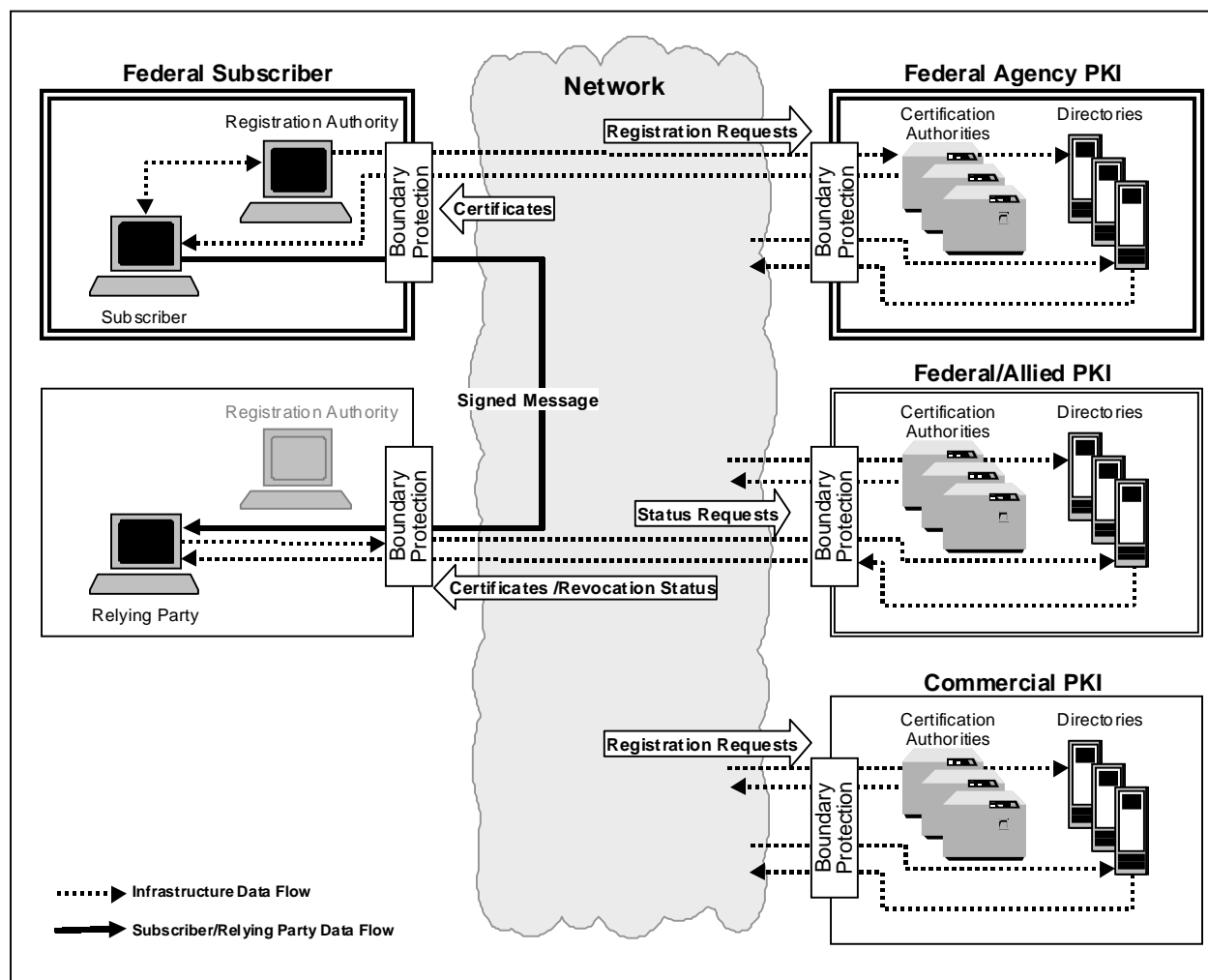
with PKIs. The section also discusses the threats and countermeasures specific to PKIs. This section is written from the perspective of the subscribers of the PKI versus that of the administrators of the PKI. The administrative perspective is contained in Section 8.1.5.12, (Administration).

### 8.1.2.1 Public Key Infrastructure Services

To support the wide variety of public key-based applications, a PKI employs a diverse set of software and hardware components, protocols, and message formats. The primary components of the PKI include Certification Authorities (CAs), Registration Authorities (RAs), and Certificate Repositories. The primary products of the PKI include asymmetric key material, certificates, and Certificate Revocation Lists (CRLs). A brief description of these components is provided below.

- **CA**— An authority trusted by one or more subscribers to create and assign certificates. [ISO9594-8] The individual operating the CA equipment is referred to as a CA operator.
- **RA** – A trusted entity responsible for performing tasks such as authenticating the identity of subscribers requesting certificates on behalf of a CA. The RA does not sign or issue certificates. RAs will generally be located at the same location as the subscribers for which they perform authentication. The individual functioning in this role is referred to as the RA operator. Many PKIs distribute the RA functions to *Local Registration Authorities* (LRAs) to provide subscribers with convenient PKI services.
- **Certificate Repository** – Location where a CA posts the certificates and CRLs that it generates so that they are available to subscribers of the PKI. Repositories can take many forms including databases and web servers, but are commonly directories that are accessible using the Lightweight Directory Access Protocol (LDAP).
- **Asymmetric Key Material** – In asymmetric or public key cryptography, two different cryptographic keys are utilized. One key is used to encrypt or sign data, while the other is used to decrypt or verify data. One key is kept secret to the entity generating the key and is referred to as the private key. The other key, which is computed from the private key using a mathematical one-way function, is made public and is referred to as the public key. Since it is mathematically infeasible to compute the private key from the public key, knowledge of the public key does not imply knowledge of the private key.
- **Certificates** – A computer based record that binds a subscriber's identity (and some authorizations) with their public key in a trust association. The certificate identifies the issuing CA, identifies its subscriber, contains the subscriber's public key and is digitally signed by the issuing CA. Often, these certificates comply with the International Telecommunications Union (ITU) X.509 standard. Such certificates are called *X.509 certificates*. [1]
- **CRL** – A list containing certificates still within their validity interval, but which no longer represent a valid binding between a public key and a particular identity. CRLs are created by a CA, and include the certificates revoked by that CA. CRLs may be posted to a repository or may be distributed through another mechanism, e.g., web and electronic mail. Other means for obtaining certificate status, such as On Line Certificate Status Protocol, are also sometimes employed instead of CRLs.

Figure 8.1-2 overlays PKI components within a generic security architecture. PKI associated with commercial entities, Federal partners, and non-Federal partners are shown along the right-side of the diagram. Note that while the figure shows Federal subscribers obtaining PKI services from Federal agency PKIs, Federal agencies will often obtain PKI services from commercial providers. Subscribers operating from secure network enclaves (normally, a Local Area Network connected to the Internet via a Firewall) work with RAs who confirm subscriber identities to obtain certificates from remote CAs.



**Figure 8.1-2 Using Public Key Infrastructures in Secure Enclaves**

Relying parties in other enclaves, associated with other PKIs, may authenticate the subscriber's public key if they trust the issuing CA. If a subscriber trusts a particular CA to correctly associate identities and public keys, then she can load that authority's public key into her cryptographic application. Any public key certificate whose signature can be verified with the public key from the "trusted" CA certificate list (trust list) and is not listed on the CRL is considered valid. This means that the subscriber's public key can be extracted from that certificate with confidence that it really belongs to the subscriber.

## UNCLASSIFIED

Key Management Infrastructure/Public Key Infrastructure (KMI/PKI)  
IATF Release 3.0—September 2000

Many CAs are required to support the validation process. Several approaches for relying parties to deal with the problem of validating their certificates issued by the numerous CAs in use are illustrated in Figure 8.1-3 and 8.1-4.

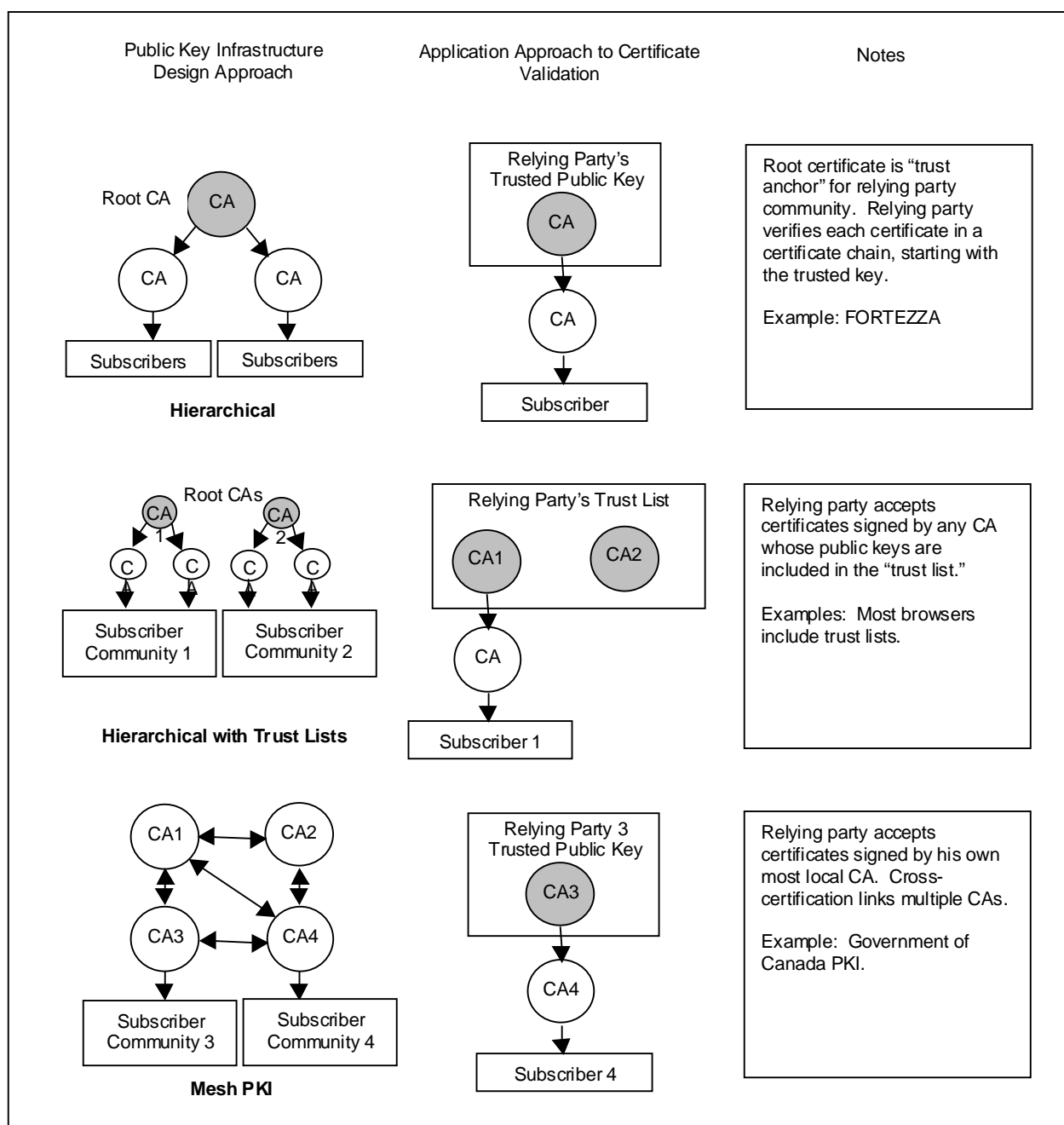
Large public key infrastructures will often support many CAs. CA may “certify” the public key certificates of other CAs. When CAs do this, they are stating that certificates issued by the certified CAs should be trusted. PKIs are often composed of a hierarchical arrangement of CAs, with a *Root CA* at the top of the hierarchy. In this way, many CAs may be certified on the basis of approval by the Root CA that serves as a “trust anchor” for the PKI relying parties.

While hierarchical PKIs have proven very popular for hierarchical organizations, many relationships within or among organizations are not hierarchical, and hence hierarchical arrangements of CAs are not always practical. For example, relying parties in the Federal Government will sometimes wish to authenticate public keys that originated from the commercial entities, academia, and foreign partners—none of whom will tolerate a subordinate hierarchical arrangement of CAs with the U.S. Federal Government.

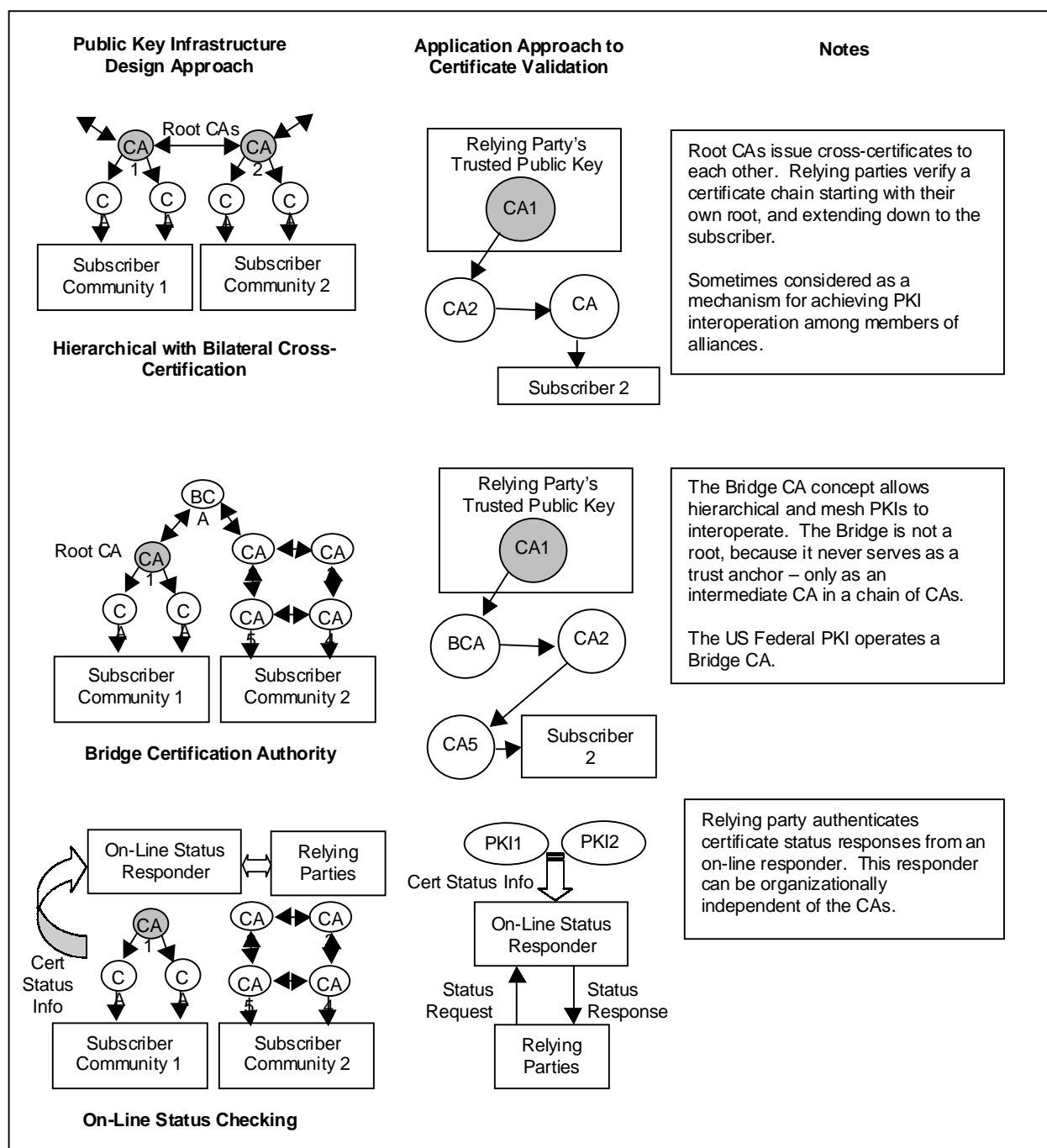
One very common way to deal with the problem of multiple non-hierarchical PKIs is to load multiple CA certificates into the verifying applications to be used as *trust anchors*. Most commercial web browsers already have over 50 “trusted certificates” pre-loaded in their trust lists by the web browser vendors. Subscribers may add other trusted certificates to this list, or delete the ones that are already there. So long as the certificate being verified was signed by a CA whose certificate is loaded in the trust list, or has a chain of certificates that terminates in a certificate included in the trust list, then the verifier considers the signer’s certificate to be valid.

Another approach to dealing with non-hierarchical PKIs is bilateral cross-certification. Bilateral cross-certification does not require a superior-inferior relationship between the CAs as is the case in hierarchical CA PKIs. Rather, two CAs—wishing to establish mutual trust among their two subscriber communities—issue certificates to each other that certify each other’s public keys. PKIs that implement such bilateral cross-certification schemes are sometimes called *mesh PKIs*, to distinguish them from hierarchical PKIs. Hierarchical and mesh PKI schemes can be combined. For example, it is possible for the Root CAs for two hierarchical PKIs to cross-certify on a peer basis.

A special case of the mesh PKI is the *Bridge CA* (BCA). A Bridge CA issues cross-certificates to Principal CAs for multiple PKIs, thus reducing the burden of bilateral cross-certification. The Federal Government is deploying a Bridge CA, which is expected to be the primary mechanism for cross-Federal PKI (FPKI) interoperation. The Federal BCA is discussed further in Section 8.1.7.4 (U.S. Federal Public Key Infrastructure).



**Figure 8.1-3 Hierarchical, Trust List and Mesh Approaches to PKI Interoperation**



**Figure 8.1-4 Bilateral Cross-Certification, Bridge CA, and on-Line Status Approaches to PKI Interoperation**

In either hierarchical or mesh PKIs, the signature verifier must build a chain of certificates that extends from the signer's public key to the CA that the signature verifier trusts. The verifier then must verify the signatures and check the revocation status for each certificate in the resulting chain. If each certificate in the chain is valid, then the verifier may consider the signer's public key to be valid.

An approach to certificate validation that breaks with the entire notion of certificate chains is that of *on-line certificate validation*. On-line certificate validation involves sending a certificate to a networked resource that has been programmed to accept or reject certificates based on the organization's validation criteria.

Each approach to achieving interoperation among multiple PKIs has advantages and disadvantages, and each has aspects that must be carefully considered if security is not to be degraded as the community of interoperation is expanded. A full discussion of these factors is beyond the scope of this document—but here are a few important points for each of the more common approaches to achieving cross-PKI interoperation:

### 8.1.2.1.1 Hierarchical CAs

#### Advantages:

- Many applications process hierarchical PKI certificates well.
- Relatively straightforward means for a large organization to enforce an organization certificate policy on a large community by revoking certificates from “subordinate” CAs not complying with the Certificate Policy.
- Application certificate processing relatively straightforward.
- Only one “Root CA” certificate needs to be distributed to the applications via “out of band” authenticated channels to provide trust in a large number of certificates issued by subordinate CAs.
- Revocation of the subordinate CAs in the hierarchy is very straightforward.
- Strong mitigation of “transitive trust” concerns – all trust decisions are made within the hierarchies of trusted PKIs (see disadvantages under “Mesh PKIs”).
- Large subscriber community can be managed using a relatively few CA certificates, providing ease of management.
- Hierarchical PKIs are usually interoperable with applications implementing trust lists.

#### Disadvantages:

- Should the Root CA certificate be compromised, the hierarchical CA's entire subscriber population is at risk, and all subscribers must load new root certificates. For this reason, Root CA keys are normally very carefully protected.
- Hierarchical arrangements of PKIs often do not parallel organizational relationships – non-hierarchical organizations (such as collections of allies, for example) often reject a hierarchical arrangement of CAs.

- PKI components based on an assumption of applications only requiring hierarchical CA elements may not be able to cross-certify, and such elements may not be able to interoperate with applications implementing mesh PKIs.

### 8.1.2.1.2 Trust Lists

#### Advantages:

- Commonly available in commercial applications.
- Relatively simple application certificate processing software.
- Provides a mechanism to provide a “per-CA” trust/don’t trust decision for each instance of deployment of a public key using application.
- No centralized management required.
- Very flexible.
- Compatible with other mechanisms of achieving trust; use of trust lists in one PKI domain does not preclude interoperation with other PKIs using other mechanisms.
- Compatible with hierarchical PKIs.
- Strong mitigation of “transitive trust” concerns—all trust decisions are made locally, or within the hierarchies of trusted PKIs (see disadvantages under “Mesh PKIs).”

#### Disadvantages:

- Management of the trust list often depends on local network administrators—or even individual relying parties—who often do not understand PKI technology, and do not have a basis for making informed decisions regarding which CAs should be trusted and which should not.
- Many applications are pre-loaded with dozens of CA certificates. Relying parties often accept all certificates issued by these CAs, without knowing anything about the level of assurance provided by the certificates these CAs issue.
- Modification of the trust list must be made relatively simple, and hence may be relatively easy to subvert (technically, or via faulty procedures).
- There is no straightforward revocation mechanism—if an organization wishes to stop trusting another CA, then the word must be spread to the organization’s relying party population, and each network administrator or individual must remove the revoked CA from all applications manually.
- PKI elements based on assumptions of trust lists may not be able to cross-certify, and applications that rely on cross-certification cannot interoperate with such PKI components.

Note that some applications allow authenticated distribution of centrally managed trust lists, which mitigate (in some cases, eliminate) many of these concerns.



### 8.1.2.1.3 Mesh PKIs

#### Advantages:

- Allows CA trust relationships to mirror business or other non-hierarchical trust relationships.
- Relieves individual subscribers and their network administrator of the burden of maintaining trust lists.
- Not susceptible to the security vulnerabilities associated with distributed management of trust lists.
- Compromise of any CA certificate affects only the subscribers of that CA; there is no “Root CA” certificate whose compromise would be catastrophic.
- Applications designed to validate mesh PKIs can usually also validate hierarchical PKIs if there is a cross certification between the mesh and the hierarchy.

#### Disadvantages:

- Developing and verifying chains of certificates from large mesh PKIs requires complex application software, and can have negative performance impacts.
- Since CAs are certifying other CAs, which may certify yet other CAs in other organizations, the arrangement of the mesh structure and the certificate security extensions must be very carefully managed to prevent the certificate chains from reflecting unintended trust relationships. This issue is sometimes called the “transitive trust” problem.
- Applications based on trust lists or hierarchical PKI concepts cannot interoperate with mesh PKIs without modification.

### 8.1.2.1.4 On-line Certificate Validation

#### Advantages:

- Simple application software.
- Relieves relying parties of the need to manage trust lists.
- Avoids the security vulnerabilities of managing trust lists.
- Avoids the management difficulties associated with mitigating transitive trust for mesh PKIs.
- Allows very rapid dissemination of revocation data. Note: Most other methods (trust lists, hierarchical and mesh PKIs, etc) use CRLs which applications pull from directory systems for revocation notification. These CRL-based revocation notification methods can be just as rapid as on-line checking depending on the frequency of CRL updates and the details of the directory implementation. On-line status checking can be seen simply as use of a special protocol for accessing a centralized trust/revocation list. The speed of revocation for such on-line methods depends on how often the centralized trust list/revocation list is updated, rather than on the speed of the on-line validation transaction. On the other hand, for large PKIs with distributed directory systems, CRL distribution and hence revocation notification can be slowed due to directory replication schemes.

- Allows applications to be compatible with all other PKI concepts, since the on-line responder can implement virtually any certificate verification technology.

#### Disadvantages:

- Requires reliable network connections between the on-line validation responder(s) and all relying parties; relying parties not able to access the responder cannot process certificates.
- Some analysts believe the centralized nature of on-line responders creates scalability problems, though such responders can be “mirrored” or replicated (perhaps at the cost of introducing the performance delays associated with directory replication).

One should bear in mind that regardless of the approach to how public key infrastructures provide for cross-domain interoperation, relying parties can only establish trust in certificates they can obtain. Many application protocols provide some or all of the signature certificates and CA certificates necessary to verify subscriber signatures, but for public key applications that encrypt data, the relying party must obtain the subscriber's encryption certificate prior to encrypting the data. This transfer of the encryption certificate (sometimes called a key management certificate or confidentiality certificate) can be accomplished via an “introductory” message between the subscriber and the relying party or the relying party can obtain the certificate from a certificate repository—often a directory system (See Section 8.1.4) Infrastructure Directory Services .

### 8.1.2.2 Security Services

The PKI plays a pivotal role in the generation, distribution, and management of the keys and certificates needed to support the public key-based security services of authentication, integrity, non-repudiation, and confidentiality. The PKI itself employs some of the security services of confidentiality and integrity. Encryption is applied to private key material that is generated, stored, and distributed by the PKI to keep the private keys confidential. Integrity services are provided to the public key material that is certified by the PKI. The digital signature on a public key certificate binds a subscriber's identity with the public key, ensuring that the integrity of the public key contained within the certificate is maintained.

### 8.1.2.3 Infrastructure Processes

There are a variety of processes or functions associated with the operation of a PKI that will be described in this section. This section is organized to reflect the KMI/PKI process categories that were described in Section 8.1.1.3 (Infrastructure Process) and summarized in Table 8.1-3. Not all of the KMI/PKI process categories apply to certificate management; processes that do not apply will be indicated.

The type of applications that PKI is supporting affects certain PKI processes. This section describes the processes in the context of two public-key based applications: secure web and secure messaging. These applications were selected due to their pervasive nature and due to the fact that they illustrate the differences between real-time (secure web) and store-and-forward (secure messaging) applications. Within this section, the differences in PKI processes that result from the influence of these different applications will be indicated. Key and certificate management for Web browsers and servers is described to show the PKI support required to enable se-

cure web communication via the Secure Sockets Layer (SSL) protocol. Key and certificate management associated with electronic mail clients is described to show the PKI support required to enable secure messaging via secure messaging protocols such as Secure/Multipurpose Internet Mail Extension (S/MIME).

After reading this section, one will note that the majority of certificate management processes are transparent to subscribers of the PKI. Subscribers only need to know the name of the CA and either its e-mail address or Universal Resource Locator (URL) in order to communicate with it. Subscriber action is required in order to generate key material and obtain certificates used within secure applications such as web and messaging. However, this interaction is part of the security configuration for a secure product and is usually accompanied by an instructive subscriber interface.

### 8.1.2.3.1 Certificate Policy Creation

A Certificate Policy states:

- The community that is to use a set of certificates,
- The applicability of those certificates (that is, the purposes for which the certificates are appropriate), and
- The common security rules that provide relying parties with a level of assurance appropriate to the community and their applications.

Before a PKI issues certificates, it should define its Certificate Policy and provide mechanisms to ensure that policy is being enforced by the PKI elements. In fact, one can consider a PKI to be nothing more than an organization's approach to generating and managing certificates in accordance with its certificate policy. This topic is discussed further in Section 8.1.5.1. (Policy Creation and Management).

### 8.1.2.3.2 Registration

The registration function is defined in Section 8.1.1.3 (Infrastructure Process) as the "authorization of people to make decisions about the validity of the subscriber actions." In general, the person responsible for decision making in the PKI context is a *Certificate Management Authority* (CMA). CMAs may be CAs (if they sign certificates or if they are responsible for a facility that automatically signs certificates) or a RA, if they simply provide the CA or CA facility with registration information. In any case, the CMA is responsible for reviewing certificate requests and verifying the information contained within the requests prior to generating certificates. The CMA operator is also responsible for authenticating the identity of the certificate requester to ensure that the proper identity is bound to the public key contained in the certificate

When an organization establishes a PKI, it will identify the personnel that will be the CA and RA operators. The qualifications such as clearances, training, etc. for the personnel who assume these roles are often outlined in the PKI's Certificate Policy (or sometimes in a Certification Practices Statement (CPS)). The CA and RA operators must also be registered with the CA or RA software being used within the system. These operators normally have special accounts that will gain them access to the administrative functions performed by the CA or RA component. To access these accounts, the operators will need to authenticate themselves to the

CA or RA components. Forms of authentication include the use of passwords, public key certificates or hardware tokens and will depend upon the capabilities of the CA or RA components used within the PKI.

While most of the security products in use today require subscriber intervention in the key generation and certificate request process, there are other models that need to be considered. One model is the case where an organization requests a set of certificates on behalf of its subscribers. In this case, the organizational representative who submits the list of subscribers requiring certificates to the PKI may need to be registered with the PKI prior to submitting the list. Registration will assure the PKI operators that the organizational request is submitted from an authorized source. Many of the CA products available today have or are adding pre-authorization features that will allow them to support this organizational registration model. Subscriber intervention is still required in the actual certificate request and response process to ensure that the proper key material and certificates are installed at the subscriber workstation.

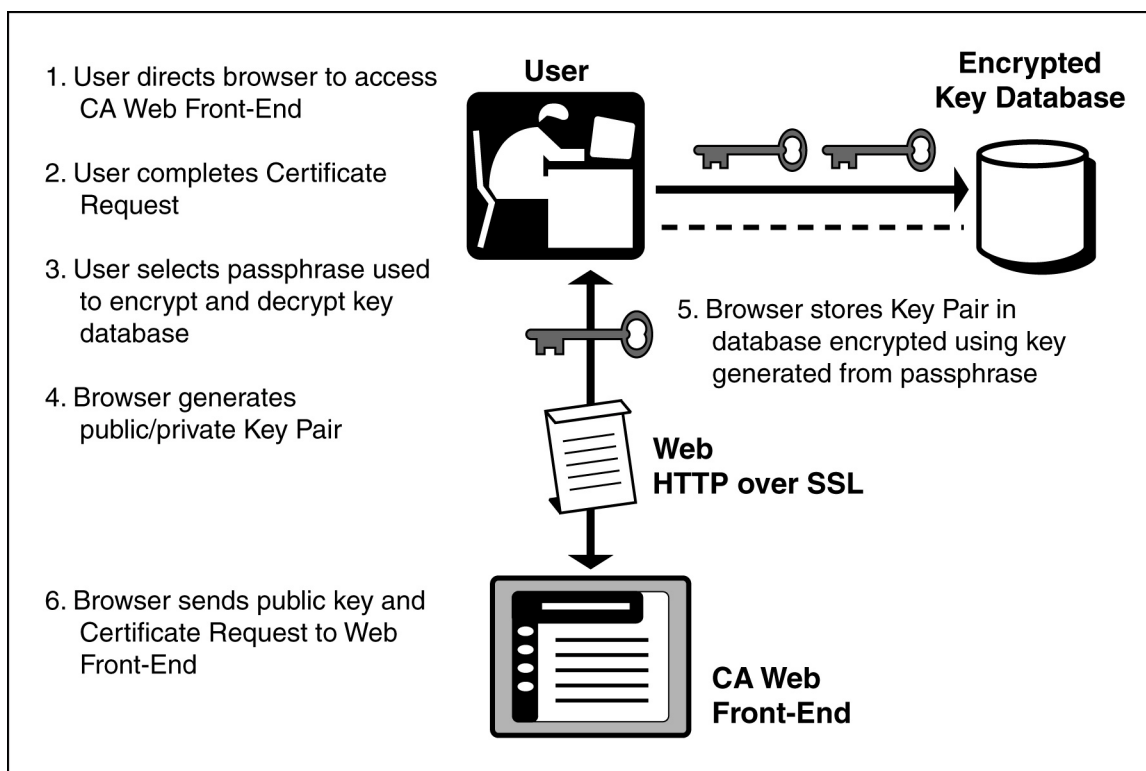
### 8.1.2.3.3 Ordering

The primary function associated with ordering in a PKI is the request for a certificate. Certain PKIs may also generate key material for a subscriber. In these PKIs, the request for a certificate will also result in the generation of a public/private key pair. Discussions of key generation by the PKI will be described in subsequent releases of the Framework in Section 8.1.2.3 (Infrastructure Processes). The remainder of this discussion assumes that the subscriber generates the public/private key pair and is indicative of the majority of secure applications in use today.

The certificate request process for Web browsers is described in detail. Differences between this process and the certificate request processes for Web servers and for S/MIME electronic mail clients will then be briefly described.

#### Web Browser

Figure 8.1-5 shows the first set of steps involved in obtaining a client certificate that is installed in a Web browser. The focus of these steps is on key generation and certificate request generation. The subscriber begins the key generation and certificate request process by directing the Web browser to connect with the CA Web front-end. The subscriber then fills in the certificate request HyperText Markup Language (HTML) form that is presented by the Web front-end. After completing the form, the subscriber presses the submit button on the form. An HTML tag (KeyGen) which appears on the form triggers the browser to generate a key pair for the subscriber. If this is the first time a subscriber has generated key material using the browser, the subscriber will be prompted to provide a pass-phrase. This passphrase is used to encrypt the subscriber's key material when it is stored in the key database that is located either on a floppy diskette or on the subscriber's workstation. When the subscriber needs to use the key material, the subscriber will be required to supply the pass-phrase, so that the material may be decrypted.

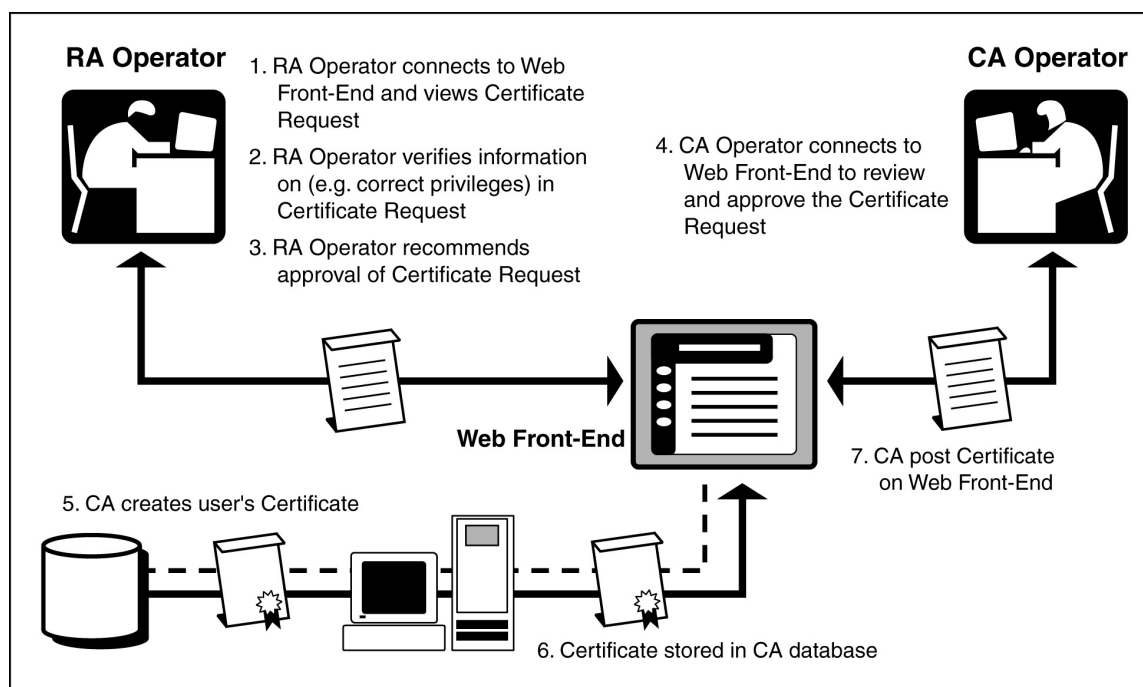


**Figure 8.1-5 Browser Certification: Key Generation and Certificate Request**

After the key material is generated, the browser provides the certificate request—which includes the public key and the information from the completed HTML form—to the server via a HyperText Transfer Protocol (HTTP) “PUT.” Most Web browsers available today support either the Public Key Cryptography Standard (PKCS) 10 [2] or Netscape proprietary certificate request format. Both of these formats are self-signed, which means that the private key corresponding to the public key contained within the request is used to digitally sign the request. The CA verifies the digital signature on the request prior to generating the certificate. This verification ensures that a private key associated with public key being certified exists and that the certificate request had not been modified in transit. Obviously, the self-signed certificate request can be spoofed. If the certificate request is captured in transit, the public key and corresponding certificate can be replaced. To counter such a threat, CAs usually only accept certificate requests across a secure channel such as a SSL encrypted session between the browser and the CA Web front-end.

The CA stores the certificate request until the RA or the CA operator approves it. Some CAs provide a reference number to the subscriber, which the subscriber can use to make inquiries regarding the status of the certificate request or to download the completed certificate.

Figure 8.1-6 shows the steps conducted by the CA to process the certificate request received from the subscriber. The certificate request approval and certificate generation process—depicted in Figure 8.1-6 and described below—assumes that the CA provides a RA function. It should be noted that the architectures of CA products vary and not all CAs have a RA component nor can they be configured to provide such a function. If there were no RA function available, then the CA Operator would conduct all steps within the certification process.



**Figure 8.1-6 Browser Certification: CA Processing Request**

The RA accesses the Web front-end to review any pending certificate requests. The RA displays the information contained in the request and verifies that it meets the policies set by the CA, e.g., if the subscriber's Distinguished Name (DN) follows the proper format or if the subscriber's key is of a certain length. If further information is required before the request can be processed, the RA can contact the subscriber who submitted the request. Other procedural activities such as requiring the subscriber to be authenticated in person by the RA may also be implemented at this point.

## Web Server

The procedure for generating a certificate for a secure Web server is similar to generating a subscriber certificate for installation into a Web browser. Most secure servers provide a forms-based interface for the Web server administrator. One of the options available through the form is to generate and install server certificates. The administrator performs the following steps for generating and installing the Web server's certificate.

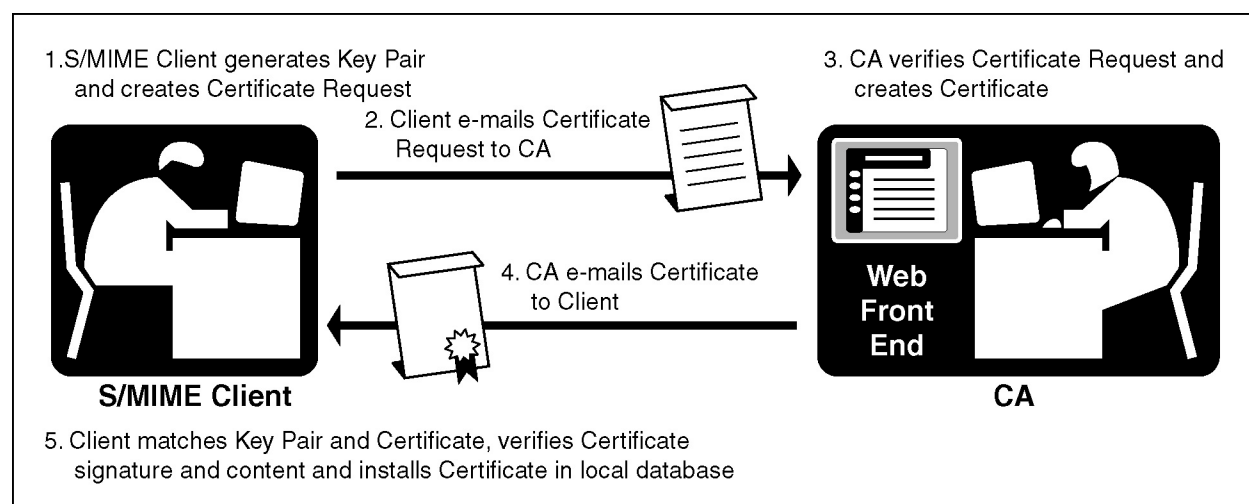
The first step is to run the key generation program at either the command line or via a Graphical User Interface (GUI). The steps for generating the public and private key pair are similar to generating a subscriber's public and private key file. The administrator must specify a file name to which the new key pair file will be stored. The administrator may need to generate random information to initialize the random number generator. Finally, the administrator must supply a passphrase that will be used to protect the key pair. After the administrator has created the server's key pair file, the administrator fills out the server's certificate request. The certificate request contains information including the server's DN and the administrator's electronic mail (e-mail) address and telephone number. Web servers utilize the PKCS 10 certificate format. After the form is completed, the administrator can send the form to the CA.

E-mail is the transport mechanism presently used by Web servers to submit certificate requests and receive certificates.

The CA process for a server certificate request is essentially the same as that of the Web browser. The only difference is the request is received at the CA via e-mail and the certificate is returned to the server via e-mail. In-person authentication of the web server is also not feasible. The CA operator can confirm information about the server request with the system administrator by requiring that the administrator appear in-person at the CA or requiring that documentation be provided by the server's owning organization which states that the server is located at that organization and requires a certificate.

## S/MIME Client Certification Process

Figure 8.1-7 shows the certification process for a generic S/MIME client. Using the security configuration options of the S/MIME client, a key pair for the subscriber is generated locally. The private key is stored in the keys database of the product. This database is protected by a key computed from hash of a pass-phrase provided by the subscriber at key generation. The public key is placed either in a self-signed certificate or in a certificate request. This description focuses on the latter option, which requires interaction with PKI components. S/MIME clients support the PKCS 10 certificate requests, which are transported to the CA via e-mail (Simple Mail Transfer Protocol (SMTP)) using a *smime.p10* message format.



**Figure 8.1-7 S/MIME Client Certification Process**

The certificate request is received at the CA via e-mail. Once the request is received, the actual generation process for the S/MIME client certificate is essentially the same as that which was followed for the Web browsers and servers previously described. In-person authentication of the subscriber may be implemented by the CA if so desired.

## 81.2.3.4 Generation

In the context of PKIs, there are two aspects of generation: key generation and certificate generation. Both generation aspects are described in this section.

## Key Generation

In public key management the generation of key material is closely tied with the request for a certificate. Therefore, Section 8.1.2.3.4 is written following a distributed model where the key material is generated locally in the context of the secure application. It is also possible to generate public key material following a centralized model where the CA or some other trusted entity would generate the key material on behalf of the subscriber or application. Since keys are generated in a single place and using only one system, centralized key generation offers the opportunity to use better equipment, e.g., cryptographic hardware, random number generators, and techniques within the key generation process. Centralized key generation is often used in environments with very strong security requirements. In addition to the location of the key generation, the models also differ in the type of additional key management functions that are required to support each model. When the key material is generated locally, the private key stays within the control of the subscriber or application from its generation to its destruction. Only the public key needs to be conveyed to the CA for inclusion in the certificate that the CA will subsequently distribute to the subscriber or application. When the key material is generated centrally, not only does the CA have to generate and distribute the certificate, but there is the added function of securely distributing the private key to the subscriber or application. Today, secure private key distribution is achieved through manual distribution or via distribution via a secure protocol, which may be either proprietary, specific to a product line, or a more widely accepted security protocol such as SSL.

Another consideration when generating key material centrally is how the key material is to be used. Usually only asymmetric key material that will be used for key or data encryption is generated centrally. Key material, which will be used for digital signature purposes, is normally generated locally. This is the preferred approach since one would like to use digital signatures to provide the security services of non-repudiation. True non-repudiation services can only be provided if the entity generating the signature key material is the only one who knows the private key. If the digital signature key material were generated centrally then this would not be the case. In light of these considerations, asymmetric cryptographic products are now migrating to two key systems where separate key material is used for data/key encryption and digital signature purposes. There are commercial products available which combine both the distributed and centralized key generation methods. These products generate key material associated with key or data encryption centrally and key material associated with digital signature locally.

Another topic associated with key generation is whether the key material is generated in software or in hardware. Many of the commercial security products available today perform all cryptographic functions including key generation in software. However, there are concerns that software cryptography may not be adequate for all situations. Therefore, there has been a move to provide flexibility within security products; to allow key material to be generated and cryptographic functions to be performed on hardware tokens including both Personal Computer (PC) Cards (a.k.a. Personal Computer Memory Card International Association (PCMCIA) Cards) and International Standards Organization (ISO) 7816 compliant smart cards. It should be noted that many of the commercial CA products available today utilize hardware tokens or other types of cryptographic hardware to generate the CA key material and perform the cryptographic functions associated with the CA functions. When hardware tokens are utilized, there are added management functions associated with the tokens themselves, including their initialization, personalization for a particular subscriber and distribution of the token and any



PIN associated with the token. Today, many of the token management functions are handled outside the context of the PKI. The FORTEZZA Certificate Management Infrastructure (CMI) is one notable exception. However, there does appear to be a trend within the PKI arena to add token management functions to the growing list of functions provided by the PKI.

Another consideration associated with key generation is the length of the key material. In general, the longer the key length the stronger the key since it is more difficult to break longer keys. In the commercial cryptographic implementations in use today, asymmetric key materials are usually 1024 bits in length, with 2048 bit or larger keys being used for more sensitive applications such as CA signing keys. Today, strong symmetric key implementations generally utilize 128 bit keys. Type 1 cryptographic implementations used to protect classified information utilize even longer key lengths. It should also be noted that export and import controls imposed by governments may restrict the key lengths within exportable or importable versions of cryptographic-based products.

## Certificate Generation

As was done in the Ordering section, a full description of the Web browser certificate generation process is provided. Differences between this process and that of the Web browser and the S/MIME electronic mail client are summarized.

### Web Browser

Figure 8.1-8 shows the steps conducted by the CA to process the certificate request received from the subscriber. If all the information within the request is satisfactory and the subscriber is authenticated to the RA's satisfaction, the RA marks the certificate for approval. Depending upon the configuration of the CA product, the certificate may be automatically generated once the RA has approved the request or CA Operator intervention may be required to generate the certificate.

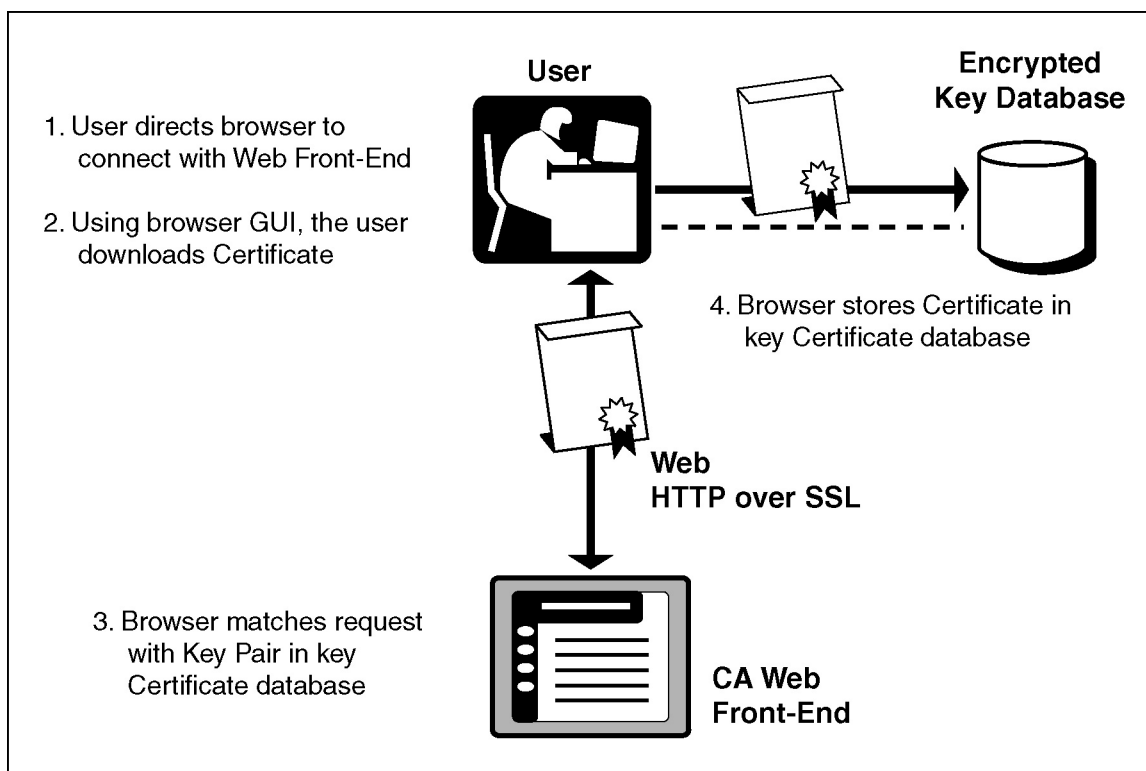
Once the certificate has been created, a copy of the signed subscriber certificate is stored in the CA database and is posted to the Web front-end. Once the certificate has been posted to the Web front-end, the subscriber can then download and subsequently use the certificate. Many CA products send e-mail to the subscriber to notify them that the certificate has been created and to provide them the URL where they may download the certificate. If the CA does not provide notification services, then the subscriber would need to periodically check the Web front-end to determine if the certificate is ready.

### Web Server

The CA process for generating a server certificate is essentially the same as that of the Web browser. The only difference is the certificate is returned to the server via e-mail.

### S/MIME Client

The certification process for a S/MIME client is shown in Figure 8.1-7. The certificate generation process for the S/MIME client certificate is essentially the same as that which was followed for the Web browsers previously described. Once the certificate request is validated, the CA generates a certificate for the S/MIME subscriber. S/MIME clients expect to receive certificates back, in PKCS 7 [3] format via e-mail.



**Figure 8.1-8 Browser Certification: Installing Certificate in Browser**

### 8.1.2.3.5 Distribution

Certificates can be distributed in several ways. The certificates can be e-mailed to the requesters or the requester can download a copy of the certificate from the Web front-end of the CA or from a Certificate Repository such as a Directory. This section describes the distribution options for certificates in the context of secure Web and messaging applications. As was done in the Ordering and Generation sections, a full description of the Web browser certificate generation process is provided. Difference between this process and that of the Web browser and the S/MIME electronic mail client are summarized.

#### Web Browser

Once the certificate has been posted to the Web front-end, the subscriber can then download and subsequently use the certificate. Many CA products send e-mail to the subscribers to notify them that the certificate has been created and to provide them the URL where they may download the certificate. If the CA does not provide notification services, then the subscriber would need to periodically check the Web front-end to determine if the certificate is ready. Figure 8.1-8 shows the final set of steps that complete the certification process.

To download the certificate, the subscriber needs to direct the browser to connect to the Web front-end. The subscriber may supply a reference number supplied during the certificate request process to find their certificate that appears as a hotlink. The subscriber clicks on this link to start the download process. Following the set of subscriber screens which the browser displays, the subscriber accepts the certificate for download. The certificate is downloaded and

stored in the keys database where it may subsequently be referenced. As part of the download process, the browser software checks that the private key associated with the public key contained in the certificate is located in the key database. If the associated key is not found in the database the software will not download the certificate and will provide an error message to the subscriber.

At certificate retrieval time, the subscriber may also need to download certificates associated with the CAs within its certification path. Usually the CA certificates are also available for download via a web interface. The download of the CA certificate is essentially the same as that of a subscriber certificate. The browser stores the CA certificate within its certificate database. However, the browser does differentiate between subscriber and CA certificates and considers the CA certificates to be trusted; meaning that certificate path validation will terminate once a CA certificate in the path is not found in the certificate database. The browser is able to identify CA certificates from subscriber certificates, since different HTML tags are applied to each type of certificates. It should be noted Web browsers are distributed with a number of well known root CA certificates (*a trust list*) already installed in the certificate database. These certificates are usually associated with vendors that provide certification services. It is possible to modify the certificate database and delete any CA certificates that one does not want to be trusted within a specific environment.

## Web Server

In general, Web server certificates are distributed to the server via an e-mail message from the CA. Certificates are often sent in a PKCS 7 [3] SignedData formatted message. This message format allows the full certification path (server and CA certificates) associated with the subscriber to be conveyed in the same message. Once the administrator receives the certificate from the CA, the administrator can install the certificate into the server. Most servers provide a GUI for this step. The GUI typically asks for the pathname to the file containing the certificate or the certificate can be pasted into a text block on an HTML form. The Web server will then automatically install the certificate in the Web server's encrypted key database. As part of the download process, the server software checks that the private key associated with the public key contained in the certificate is located in the key database. If the associated key is not found in the database the software will not download the certificate and will provide an error message to the administrator. Any CA certificates found in the PKCS 7 message will be installed within the certificate database of the Web server. Like Web browsers, the CA certificates are considered trusted and are indicated as such in the certificate database of the Web server.

## S/MIME Client

An S/MIME client receives the certificate back from the CA in an e-mail message. Like Web servers, S/MIME certificates are sent in a PKCS 7 [3] SignedData formatted message. Once received at the client, the message is opened by the subscriber. The S/MIME client provides functionality that verifies the PKCS 7 formatted message and automatically installs the client certificate and any CA certificates in the local certificate database. As with both the Web browser and server, the S/MIME client also differentiates CA certificates from subscriber certificates within its database and is normally distributed with popular root CA certificates installed. However, unlike the Web products, most S/MIME products do not automatically trust CA certificates installed in the client. Normally, the subscriber will need to explicitly mark the certificate as trusted before the client will recognize the certificate as trusted.

### 8.1.2.3.6 Compromise Recovery

This section describes how the PKI notifies its subscribers when certificates are revoked and assists its subscribers in recovering from a compromise of key material. Recovery of the PKI itself from a compromise will be described in Section 8.1.5.8 (Compromise Recovery).

There will be instances when the certificates issued by a CA need to be revoked. Revocations fall into two major categories: security compromise revocation and routine revocation. Security compromise revocation covers instances when the associated private key material has been compromised, when a subscriber no longer can gain access to the private key (e.g., forgotten PIN or password or lost token) or if the subscriber has been fired or stripped of privileges granted by an organization. Report of such compromise should be immediate and the actual revocation of the certificate by the CA should occur immediately. Routine revocation covers cases where certificates need to be revoked because information contained within the certificate is no longer valid due to a variety of reasons (e.g., name changes (marriage/divorce)) or a change of organizational affiliation. These types of revocations also need to be reported to the CA.

No matter what the reason, for compromise it is important that the CA be notified about the need for revocation. Thus, a certificate revocation notice is sent to the CA that issued the certificate. This certificate revocation notice may take many forms including an e-mail message, a phone call to the CA operator, the submission of some other type of form, or some combination of the above. It is important that the CA operator ensure that the revocation notice is authentic before revoking a certificate to prevent denial of service attacks. The request may be authenticated in a variety of ways including the use of a digitally signed revocation notice, the provision of a password or in-person authentication. Commercially available CA products are just beginning to add automated certificate revocation notification to their products so the variety of authentication options is likely to grow.

A CA notifies other subscribers when a certificate has been revoked through the issuance of CRLs. A CRL contains certificates still within their validity interval, but that no longer represent a valid binding between a public key and a DN or privilege. Certificates must remain on the CRL until their expiration date. The CA will periodically generate and distribute CRLs. CRL distribution mechanisms are usually the same as those employed for certificates; CRLs are posted to directories, made available via a Web interface or distributed via e-mail.

The distribution and process associated with a CRL is one of the major issues faced within the PKI community today. There is a concern about the timeliness of revocation notification since CRLs may only be generated on a periodic basis. To counter this, emergency CRLs or CRLs containing only certificates revoked due to compromise may be distributed on a more frequent basis and may be pushed to the subscribers versus just posted to a repository where the subscriber may need to go to retrieve the CRL. Another major concern is that the CRLs may get rather large especially as the number of certificates issued by a specific CA increases. The size of a CRL will affect the time it takes to validate a certificate path. Finally, there is often no consolidated directory from which applications can obtain CRLs. Due to these problems, many of the security products available today do not provide an ability to process CRLs or the subscribers must resort to manual methods to remove a revoked certificate from the databases of these products. At the same time, there is ongoing research exploring alternative certificate revocation models.

One such alternative is the on-line validation of a certificate. In this case a certificate or certificate path may be sent to a trusted entity—which may be a CA or a certificate repository—which will determine if the certificate(s) is valid and notify the requester of the results. On-line validation also brings its own set of concerns. On-line validation requires that there be network connectivity between the requestor and the trusted entity performing the validation. The availability of the network and the added network traffic which results from the validation requests and responses are considerations associated with implementing on-line validation. The level of trust needed in the entity performing the validation is also an issue and will depend upon the requirements of the environment in which one is operating.

A CA also assists subscribers in their recovery from a key compromise. In the case where the CA has been involved with the generation of the key material or the initialization of a token, the CA may offer backup functionality. In this case, if the subscriber has lost access to the key material and needs to recover information that may have been encrypted in that key material, the CA may be able to provide a copy of the key to the subscriber or issue a new token with the old key material provided. In the case of a security compromise, the subscriber will need to have a new key pair and certificate generated. The CA will be involved in this process to the extent it was involved in the initial key and certificate generation process that was described earlier in this section.

### 8.1.2.3.7 Accounting

A number of auditing functions associated with the PKI are described in Section 8.1.5.7 (Accounting).

### 8.1.2.3.8 Key Recovery

A PKI may provide key recovery functionality by providing key backup or escrow of key material. Key backup or escrow capabilities are normally only provided to asymmetric key material which is used to encrypt either data or keys and not to key material used for digital signature purposes. Key backup or escrow capabilities can be provided when the CA generates the keys on behalf of the subscriber. In this instance, the CA will store a copy of the private key in a secure database. This key material may be retrieved from the database and used to recover information encrypted with the material if the need arises. It is possible for a CA to provide backup capabilities even when the subscriber generates the key, but this raises the issue of how the private key is securely sent to the CA for backup. It is also possible that a completely separate infrastructure other than the PKI can be used to support key recovery.

### 8.1.2.3.9 Rekey

During the course of PKI operations, it will become necessary to renew certificates. There are two cases for renewal: one is when the certificate reaches its natural expiration date; while another is when the previous certificate has been revoked and a new certificate needs to be issued. For the first type of renewal, there are two subcategories: a renewal where both a new key pair and certificate are generated, or a renewal where the key material is not changed but a new certificate is created. Whether or not a new key pair is generated is dependent upon the recommended key life span. If the key life span and certificate validity period coincide, then new

key material should be generated at renewal. However if the key life span is longer than the certificate validity period, then it may be possible to re-certify the key material, until its recommended life span is reached.

Certificate renewal with rekey is essentially the same as the generation of an initial certificate, while the renewal without rekey may be a somewhat simpler process. CA products today vary in their renew capabilities and may limit the amount of information within the certificate that can be changed at renewal time.

### 8.1.2.3.10 Destruction

Unlike symmetric key management, the PKI is not normally involved in tracking the destruction of key material. When asymmetric key material reaches its expiration date or when it has been compromised it may be destroyed. The subscriber would normally do the actual destruction of the material. At this point in time, most security products require that a subscriber manually remove old keys and certificates from the database. It should also be noted that there are instances where a subscriber would need to retain key material even after its expiration or compromise in order to be able to recover data encrypted in this key material. In this instance, the subscriber (or an agent acting on behalf of the subscriber) will need to retain the material until access to the encrypted data is no longer needed or when the encrypted data has been re-encrypted in new key material.

### 8.1.2.3.11 Administration

Administration functions for the PKI are described in Section 8.1.5.12 (Administration).

## 8.1.2.4 Requirements

Overall security requirements for PKIs are specified in a Certificate Policy, which describes requirements imposed both on the operation of the PKI, and on PKI subscribers. General requirements for a KMI/PKI that are common to many Certificate Policies are found in Section 8.1.1.4 (Requirements). PKI subscriber requirements commonly found in Certificate Policies are described in this section. Requirements specific to the operation and maintenance of the PKI itself are described in Section 8.1.5 (Infrastructure Management). Requirements related to the use of PKI services include the following.

- Subscriber generated asymmetric key material shall be generated securely.
- The subscriber shall protect the private key material from disclosure and shall also protect any password or PIN used to access the private key material.
- A subscriber shall provide accurate information to the CA when requesting a certificate. In other words, the subscriber shall provide the appropriate identifying information and the appropriate public key for certification.
- The subscriber shall only use the private key and associated public key certificate for applications or purposes approved by the PKI. Approved applications are normally documented in the CPS for the PKI.

- The subscriber shall notify the CA when the private key has been compromised, or if other information within the certificate becomes invalid.
- The subscriber shall obtain the public key of the Root CA and any CA public key certificates from an authorized source in a secure manner.
- If an organization requests certificates on behalf of a group of subscribers, the organization's representative shall provide the CA with an accurate list of subscribers to whom certificates shall be issued.

## 8.1.2.5 Attacks and Countermeasures

The strength of the security services provided by a cryptographic capability such as digital signature depends upon a variety of factors including the security of the underlying cryptographic keys, the strength of the binding between the subscriber identity and the public key, as well as the specific application implementation. Due to its role in the generation, distribution and maintenance of private and public keys, and certificates, threats to the PKI are of concern. If the PKI operates as expected, the confidentiality of private keys and the integrity of public keys should be maintained. However, it is possible that threats to the PKI—be they intentional or unintentional—may result in the disclosure of the private keys or in the modification of the public keys. Other threats to the PKI can lead to the denial of the services provided by the system. This section focuses on the attacks and countermeasures specific to the PKI. These attacks and countermeasures are discussed from the perspective of the subscribers of a PKI. Infrastructure specific attacks and countermeasures are described in Section 8.1.5 (Infrastructure Management). More general attacks and countermeasures for KMI/PKI can be found in Section 8.1.1.5 Attacks and Countermeasures).

### 8.1.2.5.1 Attacks

Attacks aimed at the subscriber of a PKI are designed to gain access to the subscriber key material, to modify or substitute the subscriber key material, or to deny the services of the PKI to the subscriber. Attacks include the following.

- **Sabotage**—The subscriber's workstation or hardware token on which key materials and certificates are stored may be subjected to a number of sabotage attacks including vandalism, theft, hardware modification and insertion of malicious code. Most of these attacks are designed to cause denial of service. However attacks such as hardware modification and insertion of malicious code may be used to obtain copies of subscriber key material as they are generated or to obtain information entered by the subscriber such as a password.
- **Communications Disruption/Modification**—Communications between the subscribers and the PKI components could be disrupted by an attacker. The disruption could cause denial of service, but may also be used by the attacker to mount additional attacks such as the impersonation of a subscriber or the insertion of bogus information into the system.
- **Design and Implementation Flaws**—Flaws in the software or hardware that the subscriber depends upon to generate and/or store key material and certificates can result in the malfunction of the software or hardware. These malfunctions may deny services to the subscriber. The flaws may be accidentally or intentionally exploited to disclose or modify

the subscriber's key material or certificates. Improper installation of the software or hardware may also result in similar consequences.

- **Subscriber Error**—Improper use of the software or hardware associated with the subscriber's interaction with the PKI or with the storage of keys and certificate generated by the PKI may also result in denial of service, or the disclosure or modification of subscriber key material and certificates.
- **Subscriber Impersonation**—It is possible that an attacker may impersonate a legitimate subscriber of the PKI. Depending upon whether the PKI generates key material on behalf of a subscriber, the attacker may obtain both key materials and certificates in the name of the legitimate subscriber or the attacker may substitute their own key material for that of the legitimate subscriber and obtain a certificate from the PKI.

### 8.1.2.5.2 Countermeasures

Countermeasures that can prevent or limit the attacks to subscribers of a PKI include:

- **Physical Protection**—Physical protection of the subscriber's workstation, communications link with the CA, and/or hardware tokens will counter many of the sabotage and communications disruption related attacks.
- **Good Design Practices**—Concerns over flaws in the software and/or hardware design may be alleviated if good design practices are followed during the development of the software and/or hardware used in conjunction with the PKI.
- **Testing**—Testing of the software and/or hardware may also be used to counter attacks to the system that result from the exploitation of flaws in the system.
- **Training**—Training of subscribers is vital to eliminating or at least reducing the possibility of inadvertent attacks due to subscriber error.
- **Strong Authentication**—Strong authentication of the subscriber by the PKI components greatly reduces the possibility of impersonation attacks.
- **Encryption**—Encryption of the link between the subscriber and the PKI components reduces the possibility that an attacker may eavesdrop on the communications and try to disrupt or modify the communications.
- **Contingency Planning/System Backup**—Backup of a subscriber's key materials, certificates, and relevant software and hardware is the best mechanism for protecting against design flaws that result in system failure.

A Certificate Policy describes all countermeasures a PKI requires to provide a level of assurance consistent with anticipated certificate usage.

## 8.1.3 Symmetric Key Management

### 8.1.3.1 Overview

Although overshadowed by PKI in the literature, Symmetric Key Management (SKM) remains an important technique in the real world. Most legacy systems use symmetric cryptography

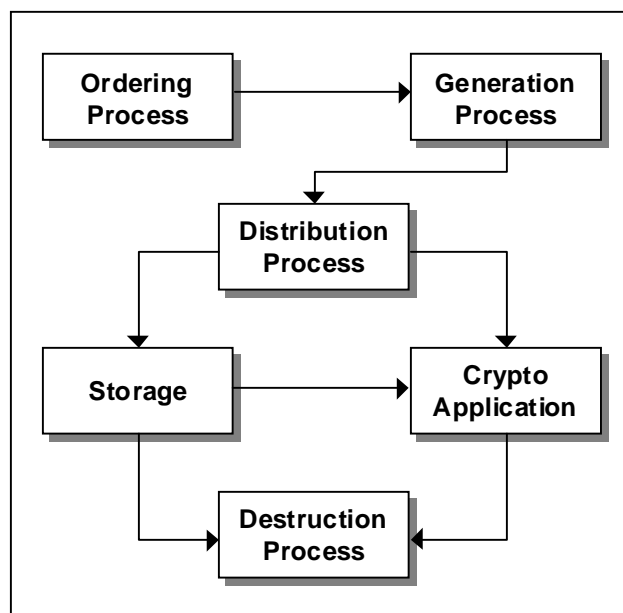


exclusively. Even with the expanding use of asymmetric techniques, many new and emerging applications, such as multicast, will still require both secure symmetric key and asymmetric cryptography.

With a symmetric key algorithm, the encryption key can be calculated from the decryption key and vice versa. This is very different from the public key algorithm where it is presumed unfeasible to calculate the decryption key from the encryption key. In most of the symmetric systems, the encryption and decryption keys are the same, requiring the sender and the receiver to agree on a key before they can pass encrypted messages back and forth. Information on certificate based public-key algorithms can be found in Section 8.1.2, Certificate Management.

The old adage “good management is the key to success” could never be more true than in the application of symmetric key in the world of cryptography. The strongest of cryptographic algorithms are reduced to nil if the management of the keys used with the cryptography is poor. For symmetric key applications where a common secret key is required by all users, delivering the correct key to all the users and keeping them secret can be extremely complicated and expensive. Figure 8.1-9 depicts the critical elements of symmetric key management.

System requirements play heavily in the decision to use symmetrical key since there are significant advantages as well as disadvantages in its use. Many of the problems with SKM have become more complex as the community of cryptographic users has increased and become more geographically separated. Ordering, generation, distribution, loading key into cryptographic applications, storage, and key destruction are becoming more critical.



**Figure 8.1-9 Critical Elements of Symmetric Key Management Activities**

### 8.1.3.2 Advantages of Symmetric Key Technology

- Everyone in a communications network can use a single key for as long as necessary. The keys can be changed as often or as infrequently as the security policy allows.
- Local generation of keys can minimize many of the problems with ordering and distribution. There is no need to connect with a central authority.
- Key structures for symmetric key is extremely simple; predominately a sequence of random numbers.

- Algorithms using symmetric key processing are generally much faster than their asymmetric counterparts. In many instances, asymmetric keys are used to securely distribute the symmetric keys to other users in the network.
- Supports both netted and point-to-point operations.
- Symmetric keying limits who holds a specific key; therefore, no outside access control mechanisms are needed to control who talks to whom.
- Symmetric keys do not require extensive validation prior to use.
- Symmetric keys are not reliant on an extended trust path.
- Potentially, fewer people need to be trusted in the ordering and distribution path.
- The creation of an unauthorized key is only dangerous when an attacker can get someone to use it in place of the correct key; consequently, by itself it does no harm.

### 8.1.3.3 Problems with Symmetric Key

- One lost key will compromise the whole network; requiring the replacement of every user key.
- Limited cryptographic services (e.g., no non-repudiation, implied authentication, etc.).
- Difficulty scaling to large communities. There is an upper limit for the size of cryptographic networks using a common key.
- The larger the number of operators using a common symmetric key, the more likely the key will be compromised.
- Large amounts of symmetric key may need to be produced to meet potential compromise and contingency uses. This key must be securely delivered and locally stored.
- Distribution delay causes key to be generated and distributed well in advance of its use; allowing potential harmful access to the key for longer periods of time.
- Nets must be predetermined. It is hard to create dynamic communication networks.
- Key must be kept secret at all times.
- Long cryptoperiods cannot be used for per session communications.
- There is no intrinsic way to know who created the key.
- No back traffic protection. A compromise of a key at any time exposes all traffic encrypted using the key since the beginning of the cryptoperiod.

### 8.1.3.4 Critical Elements of Symmetric Key Management

Good key management with its many facets is vital for maintaining security. Symmetric Key Management involves the total life expectancy of a key; controlled processes should be established and maintained for ordering, generation, distribution, storage, accounting, and destruction of the key. There must be ways to detect compromised keys and provisions to re-secure the system and efficiently determine the extent of any compromise.

## UNCLASSIFIED

Key Management Infrastructure/Public Key Infrastructure (KMI/PKI)  
IATF Release 3.0—September 2000

- **Ordering** — Only authorized individuals should be allowed to order key and only keys for which they have been given explicit authorization to order. Since the symmetric networks must be predefined, the orderer must have access to the communication network management. They need to know what users will need the key and when they will need it. The key must be ordered so that it can be delivered to all users prior to them needing it. When the key is generated centrally, it may require ordering several months in advance of actual use, given the worldwide nature of many nets. The key management system has to ensure that the orderer has the authorization to order the key as well as whether the recipient(s) are authorized to receive the key.
- **Generation** — Generation must be done in a secure environment to prevent unauthorized access to the key. The best cryptographic algorithms can be nullified if the key falls into the wrong hands. The generation process must be able to produce the total set of acceptable keys for the specified encryption algorithm. Weak or sensitive keys associated with the specified algorithm must be deleted e.g., DES has 16 weak keys. [4] Symmetric keys are usually random bit streams requiring a quality control process to insure the randomness of the bit streams.
- **Distribution** — Symmetric key can be delivered in physical form depending on trusted people and technical protection techniques like tamper resistant canisters. For very sensitive key, two-person control can be used to gain more assurance. These techniques, however, only provide minimal protection to the key over its life cycle. The more people having access to a key, the more likely it is to be compromised; therefore, a goal of secure distribution is to provide the key electronically directly from the generator to the user equipment through benign delivery techniques. Public key techniques can support benign delivery techniques. They allow the user equipment to create an authenticated session key with the generator to pass symmetric key. When true benign techniques are not possible (i.e., the user equipment does not have asymmetric cryptography), the key should be protected in encrypted channels as long as possible. Electronic deliveries to an intermediate node close to the user may be a reasonable compromise.
- **Storage** — Keys must be stored when waiting for distribution to the user or when used as contingency key. Storage of unencrypted symmetric keys may be required to recover when a link goes down. The protection of these keys is critical. They must be stored securely. Physically distributed key can only be protected through strict physical and personnel security. Electronic keys should be stored in encrypted form where physical, personnel, and computer security mechanisms are in place to limit who can decrypt and access the keys.
- **Loading Key into the Cryptographic Application** — Loading key requires a protected interface. Physical protection of the key at the interface is critical to prevent the key from being exposed where it could be copied or replaced. Although minimal protection is required for loading encrypted keys, a high level of protection is required for the less frequent loading of the corresponding protection decrypt key.
- **Destruction** — There are many potential media with which symmetric key can be deployed. These include paper (e.g., manual codebooks, key tape), mechanical components (e.g., plugs, boards), and electronic components (e.g., Random Access Memory (RAM), Electrically Erasable Programmable Read Only Memory (EEPROM), Programmable Read Only Memory (PROM)). Since the compromise characteristics of symmetric key allow recovery of previously encrypted traffic, it is imperative that the keys are not stored any longer than

necessary to perform their mission. At the end of a cryptoperiod, the secret key must be destroyed in all locations (including secondary sources like contingency storage and incidental electronic storage).

- **Compromise** — Symmetric keys are vulnerable to compromise (e.g., physical delivery, large cryptonets, long cryptoperiods), so compromise detection and recovery are critical. There are no technical mechanisms where the network can control the damage done through a compromise. The compromise of a secret key potentially exposes all the traffic it ever encrypted and invalidates the assumed authentication for future traffic. In order to recover from a compromise, each user must be notified and provided a new key. The major problems of this approach stem from the long time it might take to notify the users and then the length of time necessary to replace the keys. While users are being notified and taken off the net, other users may still be using the key thinking that it still protects the data. There are no technical mechanisms that can be used to ensure that all users have been notified. There is a significant denial of service issue bringing up a widely dispersed network. Even after a user has stopped encrypting on the compromised key, they can't communicate until the new key arrives, either from contingency stock or the generation of new key.
- **Accounting** — Due to the distribution of keys to a large number of users potentially scattered across the world and the corresponding danger of a compromised key, additional mechanisms must be in place to track keys throughout their life cycle. Effective accounting improves the tracking of who had authorized access to a key, when and where key has been delivered, and when a key has been destroyed.

### 8.1.3.5 Some Good Practices with Symmetrical Key

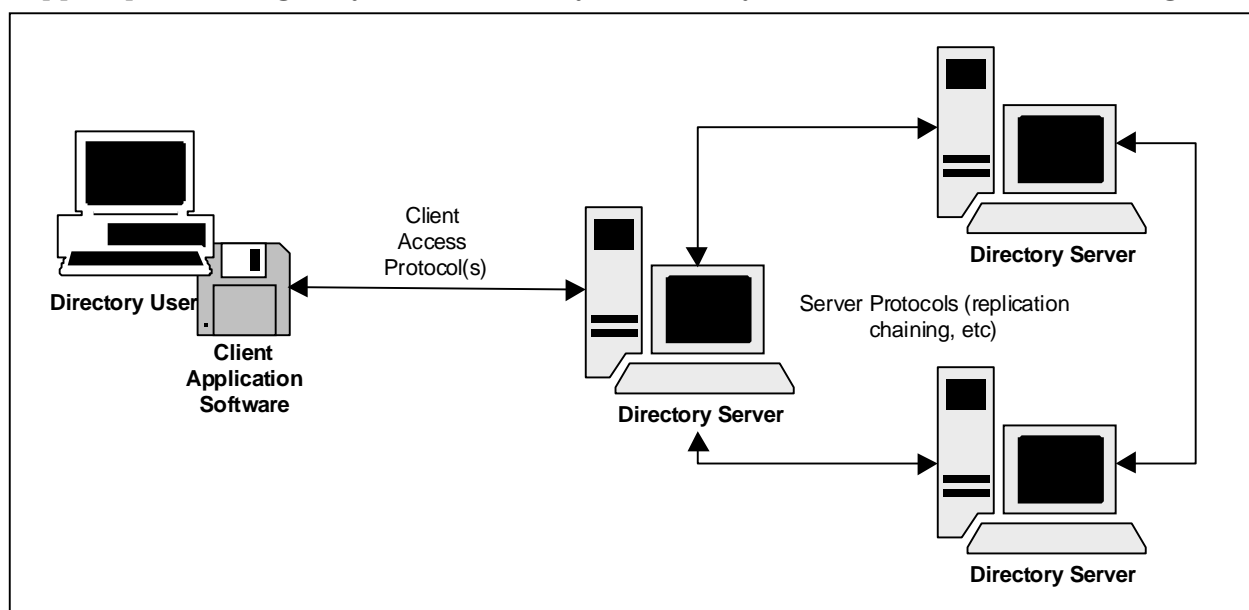
- A key order must always validate the initial requirement for the key, the number of copies, the time when the key is needed and the intended recipients.
- Revalidate requirements each time new keys are generated.
- Ensure that the person ordering and the person receiving the key are authorized for the key.
- Do not create and distribute key too early i.e., keep the storage time short. Of course, there must be enough lead time to ensure that all recipients have gotten their key.
- All key must be securely generated. This includes checks on the created key to ensure randomness.
- Secure local generation may be the best method.
- Key should be securely distributed using benign techniques where available. Where benign techniques cannot be used, limit the number of people having authorized access to the key. Use physical distribution only where absolutely necessary.
- Limit the size of the cryptonet to reduce the number of people who have access to the key.
- Limit the cryptoperiod of the key to limit the damage of an unidentified compromise.
- Limit the amount and duration of contingency key created to reduce the potential for compromise during the storage period.
- Develop procedures to quickly notify all users of a compromised key and how to replace the key with a new one.

- Train users not to use compromised key while waiting for their replacement key.
- Develop effective accounting to track the status of all keys throughout their life cycle.
- Periodically validate all key-handling procedures.
- All procedures and policies must be rigorously enforced.

## 8.1.4 Infrastructure Directory Services

### 8.1.4.1 Overview

Infrastructure directory services—through a structured naming service—provide the ability to locate and manage resources within a distributed environment. The directory also provides access control over all the objects represented within this distributed information service. Directory design can be categorized by objects within (scope of content) and functionality (range of services) supported. Within the context of this document, Directory services (See Figure 8.1-10) support provisioning of symmetric and asymmetric key material, as well as the management



**Figure 8.1-10 The Directory Model**

data for confidentiality, integrity, identification and authentication across the enterprise.

Infrastructure directory services provide a means to associate multiple elements of information with respect to a specific person or component. This association is managed in a hierarchical organization and indexed by name association. The most common example is a telephone system “white pages” that supports the association of name resolution with address and phone number elements. In the evolving distributed network environment, much more information needs to be managed, requiring more than general-purpose directory functionality. Today, the majority of deployed directory systems are considered “application-specific”, such as PKI, white pages, e-mail, or Network Operating Systems (NOS) directories.

## 8.1.4.2 Characteristics of Infrastructure Directory Services

Infrastructure directory services have several key characteristics. These characteristics are defined as:

- **Defined Name Space**—Directory services typically invoke a hierarchical namespace logically structured in an inverse tree. This naming format can be used to consolidate the accesses, easing user location of information. X.500 distinguished names, RFC822 email naming and DNS domain names may be used.
- **Highly Distributed**—Directory services reliably distribute the data to multiple servers, whether they are located across an enterprise or within a LAN environment. The mechanisms to allow partitioning of information, its access constraints and timely access are provided. Additionally, the ability to replicate data across the Directory service makes the system more resistant to failure and maintains accessibility.
- **Optimized Data Retrieval**—Directory services provide the ability to search on individual attributes of an object. The design supports a significantly higher ratio of “reads” to “write” operations. Most directory products assume 99% of the operations accessing the Directory Information Base (DIB) will be lookups and searches, as opposed to relatively few changes or additions and deletions.

Infrastructure directory services are expected to provide access to any application. Those core applications that will access directories are X.500 Directory Access Protocol (DAP); LDAP; e-mail (S/MIME V3) and a web-based access (https). Future enhancements will include support for dial-up accesses, in support of wireless key management.

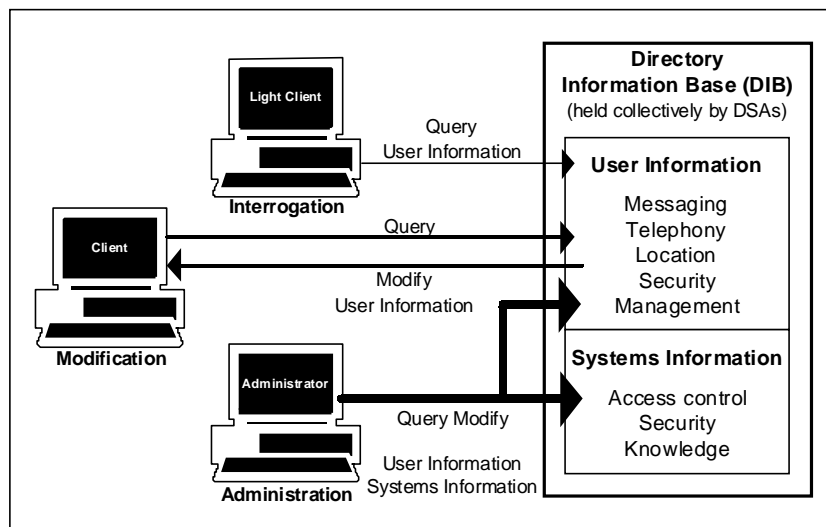
The types of clients that access directory services are:

- **Interrogation Clients**—performing general queries for user information.
- **Modification Clients**—performing queries as well as being cryptographically enabled to perform strongly authenticated binds and modification operations on selected user attributes.
- **Administrative Clients**—who have all the features of the modification client and are also permitted to manage user entries as well as operational information.

### 8.1.4.3 Information Model

The information model describes the logical structure of the DIB from the perspective of both the directory users as well as the administrators (See Figure 8.1-11). The information model defines the relationships between the objects, attributes and associated syntax in a "schema". The user information portion contains the information about a directory object that is viewable by the majority of the accesses to the DIB. The operational and administrative information portion of the DIB contains those elements of information used to keep track of directory operations. These attributes are typically schema information, access control information and information related to replicating data. Operational and administrative information is not returned in response to normal directory queries.

Further discussions related to directory distribution and Directory System Agents (DSAs) information models will be included in future releases of this document.



**Figure 8.1-11 Directory Use Access**

### 8.1.4.4 Directory Information Tree

The directory system schema is the set of rules that define how the Directory Information Tree (DIT) is constructed, defines the specific types of information held in the DIB and defines the syntax used to access the information. A schema has three components:

- **Classes**—the set of objects within,
- **Attributes** of each object class—the set of properties allowed by that class of object, and
- **Attribute syntax**—which delineates the syntactic form, as well as any matching rules used with that attribute.

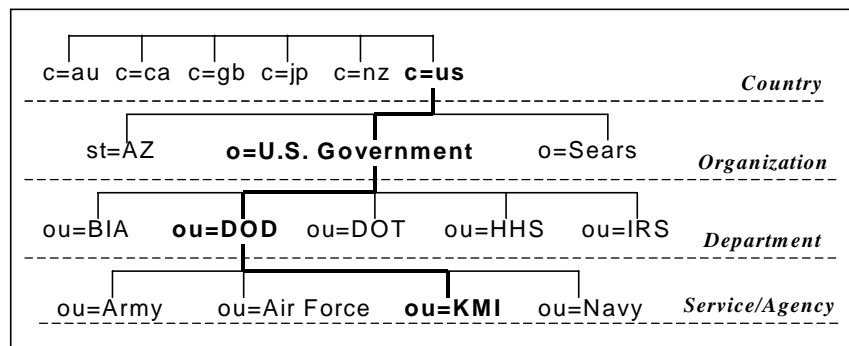
In X.500-based directory systems, an object identifier (referred to as an "oid") references object classes and attributes. In many LDAP systems, the data is essentially a string of characters, with no equivalent object identifier. This is problematic in those environments where compilers are used to interpret the data and apply cryptographic services to that data. The use of Abstract Syntax Notation (ASN.1) and associated Distinguished Encoding Rules (DER) is critical to ensuring security mechanisms applied to data in one component or domain will remain intact when utilized in another component or domain.

These three elements follow a set of rules to ensure appropriate placement of the objects into the DIT. There are content rules that identify mandatory and optional attributes within a given object class. One of the problems associated with the use of the X.509 CA object class is that it requires a *userCertificate* attribute. Thus, when the entry for the CA is created, either the CA must have the privilege to create the entry and post a certificate at the same time, or the operation will fail, violating the content rule. Many environments use directory administrators to create entries (add an object class) and allow other entities (like the CA) to populate (add) attributes at some future point in time. The newer LDAP V2 schema defines a *pkiCA* object class, where the certificate information is optional. Thus, a directory administrator can add the object class and the CAs can subsequently add the attributes with valid data.

Schema extensibility is a very useful feature to incorporate into a directory system. As new elements of data are defined, they should be added to a directory without requiring the directory to be re-started, or the compiler reconfigured. More products are gaining this feature; however, if a new object is added to the directory, consideration should be given to the upgrade of the clients that may need to retrieve and use this new element.

A DN is a sequence of naming attributes that uniquely identify an object that may be represented by an entry in the directory. Objects that may be identified using a distinguished name include organizational units, people, roles, address lists, devices, and application entities. A DN is used as the primary “key” to locate an entry in the Directory system. The DN is also typically used to identify the subject or issuer of an X.509 public key certificate.

The ***naming attributes*** which form a DN are organized in a hierarchy reflecting the DIT with a name lower in the tree identified relative to its parent entry by adding Relative Distinguished Name (RDN) attributes to the parent's DN (See Figure 8.1-12). Note that naming conventions and registration processes must be clearly articulated for a domain. Before an entry is created for an object in the directory (or a certificate created for that object), it must be allocated a DN which is unique across the enterprise. A RA normally performs the creation of a distinguished name in the directory system. Disambiguation of names is critical for key management functions; however, it is usually approached with an emotional perspective rather than a logical view. Recommendations for namespace management will appear in later versions of this Framework.



**Figure 8.1-12 Example of the Key Management Infrastructure Directory Information Tree**

## 8.1.4.5 Security Model

The security model defines the access control framework and identifies mechanisms for the access control scheme applied to a segment of the DIT. A comprehensive security model not only



addresses user access to the information within the DIB, but also includes access controls on the application itself. In addition, the security model should include the management of the cryptographic keys for identification, authentication and—if appropriate—confidentiality for the directory servers. The confidentiality services in an infrastructure directory system are typically applied at the network or transport layer.

The security services defined below are considered against the three general threats of unauthorized disclosure, unauthorized modification, and unavailability of information contained in a directory system. The information is vulnerable when held within a DSA or when transiting elements of the directory. The security services are:

- Authentication,
- Access Control,
- Confidentiality, and
- Integrity.

## Authentication

Peer entity authentication is performed between the clients and DSAs and among DSAs to provide corroboration that a user or entity in a certain instance of communication is the one claimed. The authentication mechanism can be name and password, or an exchange of cryptographically bound credentials, referred to as strong authentication. Strong authentication relies on the use of asymmetric encryption. Asymmetric encryption uses the combination of a public component and a private component to sign digitally the credentials of the user or entity authenticating itself to the system. A digital signature guarantees both the origin and the integrity of the information that is digitally signed. This binding of the public key and its holder's identification information is conveyed through an X.509 certificate that is generated by a CA. The generation of these identity certificates is usually within the bounds of an organization's certificate policy. Within a CPS should be procedures used to create, maintain, and revoke credentials for the clients, managers and directory servers themselves.

It is sound practice for all DSAs to be able to process bind requests that are name and password based, as well as strongly authenticated, utilizing an agreed upon digital signature algorithm. DSAs should support access control policy that prevents unauthorized disclosure or modification of information based on the level of authentication used. The DSA should strongly authenticate itself to its communication peer (i.e., DSAs, clients, and management entities) as required by policy. The success or failure of the steps in the authentication process should be audited and stored in the DSA audit database to facilitate compromise recovery and to enhance security of the directory.

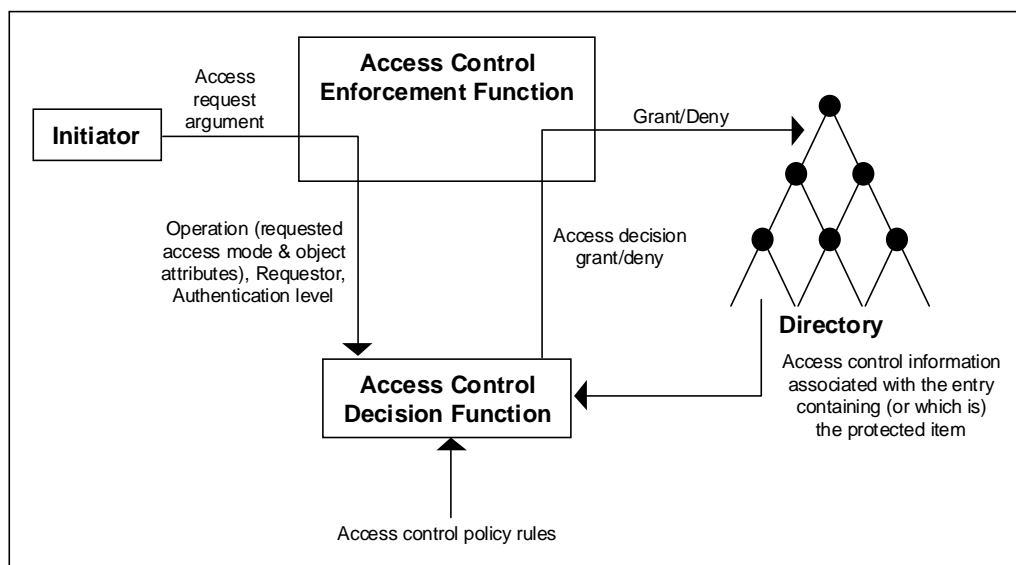
Additionally, the DSAs should not permit access to any information until all access control checks have been performed and granted. DSAs should support a standards-based (Internet Engineering Task Force (IETF), Request for Comment (RFC) 2459) signature validation process. This process should include validating the CA which produced the certificate used to sign the identification and authentication information (i.e., validate the certification path). If the path validation process cannot be completed, DSAs should reject the request and generate an audit notice. Additionally, the DSA may lock out the user from any subsequent accesses.

Once the communications partners have successfully authenticated themselves to each other, the DSA should be capable of limiting access to information stored within its DSA according to the parent (host) system security policy. The DSA should constrain setting access and privileges to authorized management entities only.

## Access Control

Access control is based on a relatively simple concept: either a list of users and the permissions to which they are entitled, or a list of protected items and the permissions necessary to access them, is held within the directory. This information is contained within Access Control Information (ACI) items. ACI items can be held within a number of parts of the directory depending on their intended usage and sphere of influence.

The Access Control Decision Function (ACDF) specifies how ACI items should be processed in order to determine whether access should be granted for a particular operation. Figure 8.1-13 is based on the ISO/IEC 10181-3 Security Framework in Open Systems standard (Part 3 - access controls). The ACDF makes the decision as to whether to grant or deny access to the requested object(s) by applying pre-defined access control policy rules to an access request.



**Figure 8.1-13 Access Control Decision Function Required for Access Control**

In some situations, the directory may not give sufficient assurance that data is kept confidential in storage, regardless of access controls. Confidentiality of attributes in storage may be provided through use of an encrypted attribute. Variations are defined in ITU-T X.501 (1997) as well as emerging IETF standards. In all instances, the directory servers do not support the encryption and decryption of this information.

## Confidentiality

Confidentiality at the application layer is an extremely difficult service to provide. It is defined in the 1997 X.500 Series of Recommendations, but relies heavily on the General Upper Layer Security (GULS) and the use of the Open Systems Interconnection (OSI) Presentation Layer. At

this point, there are no directory server products that support this service. Emerging standards permit the use of the Transport Layer Security (TLS) with LDAP, yet again, there are few, if any, products that support this service. Network and transport layer security is an extremely useful part of the layered security approach for a directory service.

## Integrity

In the event integrity is required on information stored in the directory, the information should be signed. The user who requires validation of the integrity of that information should validate the signature to ensure no unauthorized modifications have occurred. If an attribute requires integrity, the syntactical definition should expressly define it as a signed object. In the public key schema context, certificates and certificate revocation lists are signed objects.

The ability to support signed operations on all operation requests received, as well as generate signed responses to those arguments, needs to be evaluated against a performance, risk analysis and policy basis. In many cases, it is less complex and equally secure to invoke a secure channel at the network or transport layer in conjunction with the initial binding operation. Part of the security management requires the integrity protection to be negotiated and agreed-upon when establishing connectivity.

Any of the information that is stored within a Security Management Information Base (SMIB) should be protected against manipulation or destruction by unauthorized users or end entities. Changing any of the thresholds associated with collection of audit information should be made available only to authorized audit management entities. When information from one domain is replicated into another domain, the agreement to shadow should contain details on how archive of and access to audit data will be supported. Further details with respect to this critical security feature will be provided in later versions of this document.

### 8.1.4.6 Credential Management

Directory servers will require their own identity credentials when they digitally sign bind operations or other operations that may require integrity. Strong authentication is not widely deployed, but when it is, the volume of signature verifications requires either a “bank” of card readers, with duplicate hardware tokens in each reader, or some form of hardware accelerator deployed on the server hosting the directory service.

Directory Administrators (DA) will use their own sets of credentials when logging into the directory server. This permits auditing and tracking of those actions taken by the DA when modifying any of the operational information. A DSA will use their own credentials when responding to strongly authenticated bind requests, and when initiating strong binds between DSAs. In the few cases where cryptographic services are enabled in directory systems, the credentials are usually uploaded to the DSAs through a floppy interface or via a PCMCIA bus interface. The initial keying and subsequent rekeying of hardware accelerators will be discussed in future versions of this document.

## 8.1.4.7 Implementation Considerations

The directory service must have realistic performance characteristics. Performance can be seen in a number of ways: ease of use, robustness, timeliness of service restoration, and speed of access response. These aspects of the system and the generation of domain specific Concepts of Operations (CONOPS), policies, and procurement procedures will be discussed in later versions of this document.

- Ease of use is a factor of the system design and the tools presented to the directory user such as click and point, icons, windows, scripts, status messages, etc.
- Robustness deals with product and system reliability and integrity. Again, these will have to be specified in terms of Integrated Logistics Support (ILS) and Life-Cycle Costing (LCC) needs and in terms of Mean Time Between Failure (MTBF) or Mean Time To Repair (MTTR) type specifications.
- The availability goal is to provide 24 by 7 availability of any directory service. In the certificate management context, revocation information must be available upon demand.
- Service restoration deals with the recovery time for a single DSA to attain an operational state after switching on or switching the clients (and other attached DSAs) to an alternate DSA. This should not exceed five minutes if the DSA is in a strategic environment. In a tactical environment, it should be less than one minute.
- For defining the speed of response requirements, the directory system can be seen to provide two types of access characteristics. These are: the human access requirements, which deal with information retrieval (such as white pages information) via a man-machine interface or specific system functions which need to resolve for example, names to addresses for message routing. This interface is considered to be a machine-to-machine interface. Both of the above have performance requirements. However, how these are characterized and presented can be quite different. Underlying the performance of such a large-scale system is naturally the individual DSA performance and the links used between them to other DSAs and the accessing clients.

## 8.1.4.8 Client Caching Guidelines

Employing client caching is a matter of domain policy. However, the guidelines below may be followed, especially for clients caching certificate-related information.

- Store cached information in nonvolatile memory.
- Treat cached entries and cached certificates separately for the purpose of determining the useful life of the cached information. Extend the useful cache period for the certificate, since it is a relatively static entity with its own expiration time and revocation procedures.
- Capture and record, with the cached entry, the date and time that an entry was last obtained in order to determine the expiration time of that entry.
- Upon receipt of a CRL, all components containing cached certificates compare the cached certificates against the list of revoked certificates and purge those cached certificates matching the certificates listed in the CRL.
- Purge a cached certificate upon the expiration date

## 8.1.5 Infrastructure Management

The KMI/PKI infrastructure has many of the same characteristics and issues as the certificate management and symmetric key generation subscriber services described in Sections 8.1.2 (Certificate Management) and 8.1.3 (Symmetric Key Management). However, it is also a much more attractive target because a successful attack potentially subverts the security of a large number of subscribers instead of only one. In addition, it has a number of additional requirements and responsibilities not associated with subscriber services, which introduce potential new vulnerabilities. Because of these increased security concerns, the design of a KMI/PKI needs to address a wider range of issues than just supplying keys or certificates to subscribers. While the technological solutions for these problems are substantially the same as those described in Sections 8.1.2 (Certificate Management) and 8.1.3 (Symmetric Key Management) their implementation, their layering, and the procedural security solutions will be more robust. The basis of managing a secure infrastructure is trusted personnel performing their duties correctly. This section will focus on the procedural issues involved in managing the infrastructure. It will discuss unique technical requirements and issues involved with designing, developing, and operating a secure infrastructure as appropriate.

This section assumes a PKI based infrastructure with a “trusted” root element (ROOT CA) acting as a domain’s signing authority. The root element will be the basis of the domain’s trust relationship between subscribers. The root will enroll authorized infrastructure elements, e.g., subordinate CAs. These authorized elements must ensure that they only enroll other infrastructure elements that they trust. Finally, the CAs will properly identify each subscriber they enroll and ensure that their certificates are correct. The domain’s trust relationship allows subscribers to believe that the information contained in validated certificates is correct.

Building and operating an infrastructure’s trust relationship involves much more than just issuing certificates to the CAs. The KMI/PKI also has to manage itself. This requires the KMI/PKI to develop and enforce acceptable security policies and procedures, manage the key and certificate process to ensure that each element is operating correctly, manage the domain’s external relationships (e.g., determine acceptable cross-certification requirements), and ensure availability. Unique KMI/PKI management requirements include:

- Policy creation,
- Policy enforcement,
- Key and certificate accounting,
- Compliance audit,
- Cross-certification,
- Operational requirements (training, physical, personnel, operating procedures), and
- Disaster recovery mechanisms.

All PKI security attacks defined in Section 8.1.2 (Certificate Management) apply in equal measure to the infrastructure itself. However, the consequences of the attacks are now greater and it also has to protect itself against a number of new attacks that target its management of the subscribers' keys and certificates. Example attacks include:

- Global denial of service by taking down portions of the KMI/PKI,

- Substitute attacker's public/private material for KMI/PKI element's material in order to control the issuing process of subscriber's certificates,
- Destroy the domain's trust relationship via the incorporation of inappropriate elements within the KMI/PKI, e.g., inappropriate cross-certification link, and
- Compromise the data recover infrastructure.

While an attacker could theoretically attack the infrastructure to get access to an individual subscriber's information, a more likely scenario is an attacker trying to subvert the infrastructure to gain access to information on a large number of subscribers. This makes the security requirements on the internal KMI/PKI certificates stronger than on the equivalent subscriber's certificates. These increased requirements could include:

- Higher assurance in the identification process for KMI/PKI elements,
- Higher assurance in generating keys and certificates for KMI/PKI elements,
- Better protection against compromise,
- Increased mechanisms for the detection of potential compromises,
- Rigorous personnel/physical/procedural security measures,
- Stronger security architecture for limiting/monitoring operator actions, and
- Stronger data recover security.

### 8.1.5.1 Policy Creation and Management

One of the most important aspects of establishing and maintaining a trust relation for a KMI/PKI is its security policies. In order to establish the trust relationship within the domain (and other cross-certified domains), the policy must provide the basis for the subscribers to know and understand the degree of security that the KMI/PKI actually gives them. No KMI/PKI can guarantee that it is totally secure and that there is no possibility that there are unauthorized subscribers. Subscribers have to know to what degree they want to accept the KMI/PKI's assurance that the other subscriber they are communicating with is the person identified in the certificate. The only way that a subscriber can determine what trust to place in the domain is by examining the KMI/PKI's security related policy. KMI/PKIs must document their policies for both subscribers' keys and certificates and their own internal keys and certificates. Depending on the trust requirements for the specific application, these policies can range from very tight to fairly loose. Section 8.1.6 (KMI/PKI Assurance) will discuss how to define policies for applications with different levels of security requirements.

The approach to defining security policies for KMIs and PKIs tend to differ, in that PKIs are implemented in a global Federal, intra-government and commercial community, whereas KMIs tend to be operated in smaller national security communities. As a result, a great deal of effort has been devoted to developing international standards for PKI Certificate Policies, whereas KMI security policies tend to follow more local, national intergovernmental standards.

The ITU X.509 standard describes a Certificate Policy as:

“...a named set of rules that indicates the applicability of a certificate to a particular community and/or class of application with common security requirements.” [1]

An IETF informational RFC (PKIX 2527 Certification Policy/Practice Statements [5]) which defines a framework for developing policies can be found at <http://www.ietf.org/tml.charters/kix-charter.html>. The policies cover a wide range of issues from defining the rules for initializing a new infrastructure element or subscriber, to the physical and personnel requirements for the domain, to what happens in an emergency. The CP addresses issues like:

- Certification identification requirements,
- Key generation (subscriber/infrastructure, hardware/software, etc.),
- Procedural security requirements,
- Computer security requirements,
- Physical/personnel security requirements,
- Operational policy requirements,
- Requirements on subscribers (e.g., protect key),
- Interoperability requirements (e.g., cross-certification),
- Rekey mechanisms,
- Key and certificate distribution,
- Certificate profile,
- Network security requirements,
- Compromise recovery requirements,
- Liability discussion, and
- Types of applications that the certificate may be used in.

Developing Certificate Policies to the IETF Framework has proven extremely valuable in allowing an “apples to apples” comparison of PKI security practices. The IETF 2527 document has become the basis for a number of other Certificate Policy management and evaluation standards worldwide.

Certificate Policies affect the relying parties, subscribers, and those developing and deploying PKIs. They are also the basis for achieving “policy interoperability” among interoperating PKIs. For this reason, the *Certificate Policy Management Authority (PMA)*, or *Policy Authority*, needs to consider the interests of all of these parties when composing and reviewing the Certificate Policy. Furthermore, because public key certificates are often planned for use in applications having legal requirements (e.g., financial transactions) legal counsel must be an important part of most Certificate Policy development efforts.

Once created, there are a number of further actions that are necessary to make a Certificate Policy useful. The infrastructure’s approach to meeting the Certificate Policy requirements must be documented in one or more CPSs.

Certificate Policies should state high-level security requirements, and leave implementation descriptions to lower level documents, such as CPSs. In many ways, the relationship between a Certificate Policy and a CPS is analogous to that between a Request for Proposal and a Proposal. Authors of both a Certificate Policy and a Request for Proposal strive to limit their statements to functional or security requirements, and not to define specific implementations.

Authors of proposals and CPS documents strive to describe specific implementations and need to avoid simply repeating requirements. The PMA is responsible for reviewing CPS documents to ensure they meet the PKI's Certificate Policy requirements.

The CPS documents should be distributed to the PKI elements responsible for fielding and operating the PKI. The KMI/PKI components are procured or designed to the specifications of the approved CPS implementation document, and personnel are trained in the procedures defined in the CPS. During operation, the KMI/PKI must employ mechanisms to enforce—and document—that the CPS provisions are correctly followed by the PKI. Usually, such enforcement consists of a regime of compliance audits conducted by third-party auditors (or other professionals).

Finally, the policies should be periodically reviewed, updated, and distributed to ensure that they still provide the necessary security. Without these actions, the subscribers have no idea how much trust to place in a key or certificate.

Attacks against the policy creation process can disrupt the domain's trust model by misrepresenting the level of security provided by the KMI/PKI. While this misrepresentation does not lead to any direct attacks against either the KMI/PKI or the subscriber data, it may permit the key or certificate to be used in inappropriate applications where other attacks may be successful.

## 8.1.5.2 Registration

Subscribers typically “trust” the local element that provides their key or certificate because in a normal office environment, the local operator is often someone known to the individual. The subscriber also generically “trusts” the KMI/PKI root, which might be the company personnel office. The KMI/PKI trust relationship relies on the fact that every other infrastructure element—and by inference every other subscriber—is just as reliable as those elements which the subscriber personally trusts. Cross-certification extends the trust relationship to all infrastructure elements in all the other cross-certified domains.

The abilities to approve new CAs and to cross-certify other domains are critical functions that must be strictly limited. Registration is the procedural process for identifying to the infrastructure the people and elements authorized to change the domain's trust relationship. For infrastructure elements, there are normally two separate processes involved. The first reviews the policy implications of adding a new infrastructure element or allowing cross-certification. This is a procedural process done out-of-band by the Certificate Policy Management Authority. The second process is to implement the policy decision by creating the appropriate certificates. The persons responsible for implementing the decision are the root and CA operators.

When a domain establishes an infrastructure, it will identify the root and CA operators. The Certification Practice Statement (cf. Section 8.1.5.1 (Policy Creation and Management)) should outline the qualifications, such as clearances and training, for the personnel who assume these roles. The operators must also be registered with the software being used within the system. These operators normally have special accounts for access to the administrative functions at each component. To access these accounts, the operators will need to authenticate themselves to the components through the use of passwords, public key certificates or hardware tokens. The components need to ensure that these authentication processes are strong enough so that



an attacker cannot gain access to these special functions. The effect would be that the attacker could enroll an infrastructure and hence unauthorized subscribers.

### 8.1.5.3 Ordering and Validation

The ordering process within the infrastructure consists of two phases; making a request to the registered authority to add a new infrastructure element or cross-certification, and providing the necessary information to generate the certificate, e.g., CA's identity, CA's public key, in a secure, authenticated manner. The ordering process validates the request and provides a mechanism for protecting the integrity of the public key and authentication information. The generation process will bind the authentication information into the certificate.

While the electronic ordering mechanisms discussed in Section 8.1.2.3 (Infrastructure Processes) can establish new KMI/PKI elements, due to the sensitivity, an off-line manual process is more likely. Complicating the issue is the possibility that in many domains, the new element will not be in close physical proximity with its superior element. In this situation, the enrolling CA will not be able to personally identify the ordering CA.

While subscriber's orders only require validation of their identity and the correctness of their certificate information, an infrastructure element must show that they properly implement the domain's policy. This requires that before the KMI/PKI generates a certificate for an infrastructure element: 1) it establishes the need for the new element with its specific set of privileges, 2) the element understands the policy and complies with its requirements, and 3) the set of people who are operating the element are trustworthy.

Cross-certification is also likely to be an off-line manual process. However, it is likely that the two domains will not be in close physical proximity and will not be able to rely on personal identification. Prior to generating a cross-certification certificate, the KMI/PKI must validate the request. Beyond establishing the identity of the domain and its certificate information, this requires that the KMI/PKI establish the need for a cross-certification with this particular domain, determine that the policies of the two domains are consistent, and ensure that the other domain complies with its stated policy.

### 8.1.5.4 Key Generation

Please refer to Sections 8.1.2.3 (Infrastructure Processes) and 8.1.3 (Symmetric Key Management). There are no unique infrastructure requirements. Given the additional threat against the infrastructure, it needs a higher degree of assurance in the keys. This can take the form of longer keys, hardware key generation and storage, or input from multiple elements.

### 8.1.5.5 Certificate Generation

Once a new CA is authorized, the technical process of creating and signing a certificate for the infrastructure and the subscribers is similar to process for subscriber certificates (cf. Section 8.1.2.3 (Infrastructure Processes)). The primary difference is that the infrastructure must generate the initial root key and certificate in a unique way. Certificates for the other KMI/PKI elements and subscriber are identical. There will also be some differences in the certificate's

profile; however, because some of the X.509 v3 certificate fields only apply to the infrastructure and some only apply to the subscribers.

The root certificate is unique because it is self-signed, so there is no higher level device that can generate the certificate. This creates a unique process in a security critical function. The root performs the following activities to initialize the domain.

- Create the domain's cryptographic parameters (when required).
- Output the domain's cryptographic parameters in order to distribute them to the subscribers.
- Generate a public/private signature key.
- Generate a root certificate signed with the private signature key.
- Output the public signature key in order to distribute it to the subscribers.

The biggest difference in the certificates is that infrastructure certificates populate the constraint and policy fields to limit the ability of a compromised KMI/PKI element to affect other elements. The generation process has to ensure that the certificates are appropriate for the certificate's application. The specific fields populated depend on the domain's policy. The Federal PKI certificate profile, which can be found at <http://csrc.nist.gov/pki>, identifies the following profile: [6]

The certificate profile identifies four types of certificates with different requirements: root, general CA, cross-certificate, and end subscriber. All types of certificates use the complete set of X.509 base certificate fields except issuerUniqueIdentifier and subjectUniqueIdentifier. It is in the extension fields that the various certificates differ. The root certificate only populates two extensions, subjectKeyIdentifier, which identifies the specific root key being used, and basicConstraints, which identify it as a CA. The CA and cross-certification certificates are similar. They must process (although not necessarily use) all extensions except privateKeyUsagePeriod, and subjectDirectoryAttributes. Three fields—policyMapping, nameConstraints, and policyConstraints—used in infrastructure certificates are not used in subscriber certificates. The profile identifies other differences in the specific fields for each extension.

The root private key is the most valuable key in the domain. If compromised, the attacker can create unauthorized certificates that allow him to masquerade as anyone in the domain. Since the root certificate is self-signed, it is uniquely vulnerable to substitution attacks. If an attacker can get a subscriber to believe that their self-signed certificate is from the root, then they can issue certificates that the subscriber will believe are valid. Also, if an attacker can force the root to use a known key or generate a key susceptible to cryptographic attack, then they can generate their own root certificate. Also, it is likely that there will be a stored copy of the signature key in case of a root failure. If the root fails and there is no signature backup, the entire domain has to be reinitialized with the new root certificate. These security issues highlight the extreme care that the infrastructure must take to protect the root key and any copies that might exist.

## 8.1.5.6 Distribution

The KMI/PKI has to ensure that all subscribers in the domain have authenticated access to the necessary system information and certificates. The directory discussed in Section 8.1.4

(Infrastructure Directory Services), will be one method of distribution of certificates and other parameters. The infrastructure has to distribute four items: the system parameters, its own certificates, compromise recovery data, and subscriber certificates.

The authenticated delivery of the system parameters, including the domain's cryptographic parameters (when available) and the root certificate, are security critical because they are the foundation of the domain's trust relationship. While they are public values, their authenticity is critical to the correctness of the subscriber's certificate validation process. The parameters, created by the root during system initialization, are used by the CAs during the generation of other certificates. Distribution mechanisms may include a directory, off-line distribution, or local distribution through the CA. The KMI/PKI must also ensure that all subscribers have authenticated access to its certificates and compromise recovery information.

After certificate generation, the KMI/PKI provides the subscriber with their certificates. Prior to activating a new certificate, both the infrastructure and the subscriber should check that the certificate was properly generated. The infrastructure has to check that the certificate owner has access to the private key that corresponds to the certificate's public key. Proof of Possession (POP) is one protocol solution for performing this check. The subscriber has to check that the certificate contains the correct public key and subscriber information. After completing the checks, the subscriber indicates to the infrastructure to post their certificate.

## 8.1.5.7 Accounting

The KMI/PKI has to be able to track the location and status of keys and certificates throughout their life cycle. There will likely be a requirement to archive the accounting information because of the legal need to be able to document the life history of a key or certificate for as long as the signature might need to be verified. The accounting information for each certificate should provide, at a minimum, the certificate contents plus the applicable information for each task, including the following.

- Task.
- Time.
- Status (completed/error).
- Operator involved.
- Element that originated the task, e.g., where did the order originate.
- Other element(s) involved in the task.
- Acknowledgment from other element(s) involved.

Accounting has both real-time security and administrative requirements. It provides a security service by allowing the check that each step of the process was proper (e.g., the certificate generation process checks the status of the order validation) prior to the beginning of the next task. Accounting also tracks the interaction between various components by requiring each element to acknowledge to other involved elements that it has completed its portion of the processing.

The primary use of an account is administrative. The system needs to be able to track the history of keys and certificates in case of future challenges to its authenticity. Accounting is useful for the following tasks.

- Showing an outside observer the infrastructure life cycle for any key.
- Proving to an outside auditor that the policies and procedures were correctly followed.
- Providing damage assessment of operator actions if an operator is subsequently shown to be untrustworthy.
- Recording certificate information from the ordering process.
- Archiving a key's history.
- Archiving a token's history.

Depending on the KMI/PKI architecture, a single element or many elements can perform the accounting. All accounting records must be protected against accidental deletion or modification, or malicious attacks. If several elements perform accounting, either for one key or certificate or because multiple certificates from different elements reside on one token, there is an additional issue of coordinating the partial accounting records into a complete, authenticated set of records.

## 8.1.5.8 Compromise Recovery

An infrastructure element can compromise either its signature key or key agreement key. The compromise of a KMI/PKI element's key agreement key is the same as for a subscriber's key (cf. Section 8.1.2.3 (Infrastructure Processes)).

Since the compromise of an infrastructure element's signature key invalidates all lower level certificates that include the element in their validation path, it is the more serious problem. This includes the direct certificates it created for lower level CAs and subscribers as well as any certificates created by the CAs. It is critical that the infrastructure must be able to reenroll the affected elements and subscribers quickly and painlessly, while removing any unauthorized subscribers enrolled by the compromised element. The infrastructure must be able to inform the subscribers and cross-certified domains about an infrastructure compromise quickly and accurately, while rapidly rekeying the affected elements and subscribers. The responsibility for informing the subscribers resides in the element that enrolled the compromised element. The mechanisms for notifying subscribers about the compromise of an infrastructure certificate are the same as those defined in Section 8.1.2.3 (Infrastructure Processes) a CRL or on-line verification.

For a compromised root, the same mechanisms theoretically work, but it is unclear that the applications support will be there. Possible solutions include placing the root certificate on a root generated CRL, placing the root certificate on the PCA CRL, or on-line verification. When checking a CRL, normal processing is to look for the certification on the CRL from the enrolling CA. Both possible CRLs for the root (its own or from a subordinate CA) are exceptions to this processing and it is unclear if the applications will support them. On-line verification protocols are still in the design stage and it is unclear if they will report the root as compromised. Alternative workarounds, such as placing every CA on the appropriate CRL, may meet the requirement.

The recovery process for re-enrolling subscribers is straightforward, but must be performed quickly to minimize the impact on the subscribers. Starting at the compromised element, it generates a new public/private key pair and a higher level element generates and signs and

distributes the new certificate. Once the element is operational again, it can begin to reenroll its subscribers. The re-enrollment process requires a revalidation of every subscriber, using any of the mechanisms outlined in Section 8.1.2.3 (Infrastructure Processes). An issue is how to deal with the occasional PKI subscriber that has not tried to validate a certificate since the compromise. They will not realize they need to be re-enrolled. The infrastructure can allow them to continue to have an unusable certificate or it can contract them about being re-enrolled. Lists of subscribers should be available from either the local accounting records or the directory.

### 8.1.5.9 Rekey

An infrastructure element's rekey process differs for key exchange key and signature key. An infrastructure element's key exchange key is similar to a subscriber's rekey addressed in Section 8.1.2.3 (Infrastructure Processes). Signature rekey has major impacts on the CAs or subscribers created by the element, so the KMI/PKI has to give strong consideration to how often they rekey the infrastructure elements. The consequence of rekeying an infrastructure element's signature key is that every certificate in its verification chain must also be re-keyed. This creates a tradeoff between security and subscriber friendliness over the frequency of rekey. Security considerations push for frequent re-keys because of the consequence of an undetected compromise or a crypt-analytic attack of an infrastructure element. Subscriber friendliness demands infrequent re-keys because of the impact on the subscribers of rekeying the infrastructure.

The security tradeoffs are straightforward. The private signature key of an infrastructure element is a high value target because a compromise allows an adversary to masquerade as anyone in that element's domain. The longer the key remains in use, the greater the incentive for attacking it, and the better chance the adversary has of being successful. Once the element is re-keyed, the old signature key has no value.

Infrastructure rekey operational issues that should be included in the process are listed below.

- There should be a graceful rollover to the use of new keys without a period of community isolation or non-interoperability.
- Revocation notification must be maintained during the rollover. This means that KMI/PKIs will probably maintain multiple, simultaneously current CRLs.
- Note that a CMA may continue to sign CRLs with the old key, long after it has ceased signing certificates with that key and until the last certificate signed with that key expires and its CRL-inclusion period passes.
- It should be possible to issue certificates that will not fail validation due to expired signing authority certificate, i.e., the requested certificate should verify for a reasonable period of time even when issued just prior to rekey of the signing authority. (This is often accomplished by making the signing authority certificate validity period longer than the signing authority private key usage period.)
- The issuance of certificates should not be unreasonably delayed when authority rekey is pending. That is, an end subscriber certificate request should not kick off an authority rekey, possibly extending to multiple levels of the hierarchy, for which the subscriber must wait.

- The mechanism will have to live within the constraints of the cryptographic token(s) employed at the time of its introduction.

One method of minimizing the subscriber impact is to use the current key to authenticate the new key. The steps to initiate this action are listed below.

- The Root CA generates a new key and creates a new certificate with its public key signed with the current signature key.
- The Root CA also creates a new certificate for the current key and signs it with the new signature key.
- Subscribers needing the old CA certificate containing the old key must cache it locally, as it will not be available from the directory.
- All subordinate subscribers and authorities should be notified of the impending rekey, so that they can cache the certificate containing the old key and, probably, the last CRL signed by the old key.
- Applications must recognize when data is signed using a private key associated with an old certificate and obtain the old certificate from its cache.
- Applications may have to forego checking of current CRLs issued by the rekeyed authority and incur the associated risk.
- All subscribers and authorities whose certificates were signed by a re-keyed authority should obtain new certificates, signed by the new key, as quickly as possible.

CAs will continue to issue CRLs signed by the old key until one CRL-inclusion period after the expiration of all certificates they have issued. Therefore, subscribers can continue to be notified of revocations of certificates signed by the old key.

When it is time for the Root CA to rekey, the subscriber can validate the signature regardless of which key the sender and recipient have. For example, if the Root CA and the sender have both been re-keyed but the recipient hasn't, the recipient's validation chain would be as follows: the sender, its CA(s), the new Root CA certificate, and finally the new Root CA certificate signed with the old signature key. Once the Root CA begins its rekey process, each of the CAs can use a similar process to generate their new keys.

## 8.1.5.10 Destruction

Please refer to Section 8.1.2.3 (Infrastructure Processes).

## 8.1.5.11 Key Recovery

There are two separate issues about key recovery in the infrastructure. The first deals with how KMI/PKI elements perform key recovery. The second deals with the issues involved in developing a key recovery infrastructure.

## Key Recovery for KMI/PKI Elements

There are no easy answers about the requirement for key recovery in infrastructure elements. It depends on the policy of the domain. This section defines some of the tradeoffs in the key recovery policy.

In general, signature keys do not need key recovery. The signature key serves no law enforcement purpose and the subscriber suffers no great inconvenience in getting a replacement signature key. Within the infrastructure, however, the enormous impact of rekeying the element and its subscribers for lost or destroyed keys (cf. Section 8.1.5.9, Rekey) drives the requirement for key recovery of certain signature keys. The policy can limit key recovery to only certain elements. Even if some elements, such as the root, require key recovery, other elements within the infrastructure do not. Given the obvious security ramifications of storing signature keys, a robust recovery system must be in place to protect keys against all unauthorized access. The key recovery policy for KMI/PKI element's key agreement keys is the same as for any other domain subscribers.

## Key Recovery Infrastructure

There is no one key recovery infrastructure. Either the certificate management infrastructure or a completely separate infrastructure can perform key recovery. The regular certificate management infrastructure would store encrypted keys at the CAs. Advantages include no additional people with access to the key, lower cost, and infrastructure employees could already exploit the keys through other attacks. A separate infrastructure could use any approved method. Advantages include potentially tighter security for the keys and no political fallout for the certification management infrastructure. The next section describes a generalized recovery architecture based on the draft Key Recovery Federal Information Processing Standard (FIPS).

## Generalized Key Recovery Model

The key recovery system model defines the minimal set of system components needed to perform key recovery. The key recovery system model is a generalized model that supports a wide variety of different key recovery techniques and data applications. The key recovery system model contains the following components, as a minimum.

- System A (Encryption-Enabled).
- System B (Encryption-Enabled).
- Recovery Information Medium (RIM).
- Key Recovery Requester System (Requester System).
- Key Recovery Agent(s) (KRA).

The model depicts a key recovery system capable of creating key recovery information (KRI) and recovering the key from the key recovery information.

The three components—System A, System B, and the key recovery information medium—collectively define the “Key Recovery Enablement Process.” The process also includes an Encrypted Data Medium and a Key Distribution Medium. The Encrypted Data Medium and

Key Distribution Medium are the “locations” where the encrypted data and data encryption key are stored or communicated, respectively.

The process of encrypting data and creating key recovery information is divided between one or more encryption-enabled systems, denoted in the key recovery system model as *System A* and *System B*. An *encryption-enabled* system can encrypt and decrypt data. System A, System B, or both need the ability to create key recovery information. However, the key recovery system model does not prescribe which system or systems must have a key recovery capability. The *RIM* maintains the key recovery information produced by these systems. The *RIM* may exist over multiple “locations”, and may be in the same or different location from the Encrypted Data and Key Distribution Mediums.

The *RIM* represents the “locations” where the key recovery information is stored or communicated, such as a storage device or a communications channel. The key recovery system model does not prescribe how or where the key recovery information must be stored or communicated, so long as the *RIM* is available. In order to allow interoperability between various key recovery schemes, a standard format for key recovery information on the *RIM* is essential. Each scheme has a distinct set of information that must be present in order to allow key recovery. A *Key Recovery Field* (KRF) contains this information. To ensure the integrity of the KRF, the association of the KRF with the encrypted data, and to provide the identities of the key recovery scheme in use and the appropriate KRAs, a *Key Recovery Block* (KRB) contains the KRF.

The key recovery information itself is managed or handled in a variety of ways. It may exist for only a brief time during electronic transmission, or it may exist for a relatively long time on a storage device.

The Requester System and the KRA form another sub-portion of the key recovery system model called the *Key Recovery Process*.

The *Requester System* and *KRAs* handle the process of recovering a key from the key recovery information. They access the Encrypted Data Medium and the *RIM* and interact with one or more *KRAs* using a Requester System to recover a cryptographic key from the key recovery information.

A recovered key can then be used to recover the data, either directly or indirectly, using a Data Recovery System. The data encrypting key is recovered *directly* when the recovered key is the same key used to encrypt the data. *Indirect* key recovery occurs when the recovered key is a key encrypting key used to decrypt or recover the data encrypting key.

## Requirements

This section defines some of the security requirements on a key recovery infrastructure and its elements. It discusses a high assurance commercial level recovery infrastructure. Depending on the application, higher or lower assurance infrastructure may be appropriate.

### Key Recovery Agent Requirements

- Cryptographic Functions—All cryptographic modules shall be FIPS 140-1, Level 3 compliant.



## UNCLASSIFIED

Key Management Infrastructure/Public Key Infrastructure (KMI/PKI)  
IATF Release 3.0—September 2000

- Cryptographic Algorithms—The key recovery scheme shall be at least based on using only FIPS algorithms. The implementation of these algorithms shall conform to the applicable FIPS standard(s) (Same as Level 1).
- Confidentiality—
  - The KRA shall protect key recovery information stored against disclosure to unauthorized individuals.
  - The KRA shall protect key recovery information transmitted against disclosure to parties other than the requestor(s).
  - The KRA shall prevent any single subscriber or mechanism from compromising the confidentiality of the key recovery information.
- Audit—
  - The product/system shall be able to generate an audit record of the following auditable events.
    - Start-up and shutdown of the audit functions; and
    - All auditable events as defined in the system security policy.
  - Examples of auditable events include the following.
    - All requests to access subscriber authentication data.
    - Any use of the authentication mechanism. The authentication information shall not be stored in the audit trail.
    - All attempts to use the subscriber identification mechanism, including the subscriber identity provided.
    - The addition or deletion of a subscriber to/from a security administrative role.
    - Requests, responses, and other transactions generated by the product/system.
    - Requests, responses, and other transactions received by the product/system.
    - The invocation of the non-repudiation service.
  - The audit event shall include identification of the information, the destination, and a copy of the evidence provided. The event shall exclude all private and secret keys in encrypted or unencrypted form.
  - The product/system shall be able to associate any auditable event with the identity of the subscriber that caused the event.
  - The product/system shall be able to generate a human understandable presentation of any audit data stored in the permanent audit trail.
  - The product/system shall restrict access to the audit trail to the authorized administrator.
- Identification and Authentication (I&A)
  - The product/system shall provide functions for initializing and modifying subscriber authentication data.
  - The product/system shall restrict the use of these functions on the subscriber authentication data for any subscriber to the authorized administrator.
  - The product/system shall protect from unauthorized observation, modification, and destruction, authentication data that is stored in the product/system.
  - The product/system shall be able to terminate the subscriber session establishment process and disable the subscriber account after five unsuccessful authentication attempts until an authorized administrator enables account.

## UNCLASSIFIED

Key Management Infrastructure/Public Key Infrastructure (KMI/PKI)  
IATF Release 3.0—September 2000

- The product/system shall authenticate any subscriber's claimed identity prior to performing any functions for the subscriber.
- Access Control
  - The product/system shall verify applicable authentication and integrity services for the received transactions as determined by the standard compliant protocol.
  - The product/system shall apply applicable authentication, integrity, and confidentiality services to all transactions, i.e., requests and responses, as determined by the standard compliant protocol.
  - The product/system shall release the keys only to authorized subscribers.
  - The KRA shall release the key only if the requester is authorized to receive the key associated with the subscriber specified in the request and for the validity period (time) if specified in the request.
  - The product/system shall ensure that security policy enforcement functions are invoked and succeed before any security-related operation is allowed to proceed.
  - The product/system shall restrict the ability to perform security-relevant administrative functions to a security administrative role that has a specific set of authorized functions and responsibilities.
  - The set of security-relevant administrative functions shall include all functions necessary to install, configure, and manage the product/system. Minimally, this set shall include assignment/deletion of authorized subscribers from security administrative roles, association of security-relevant administrative commands with security administrative roles, assignment/deletion of subjects whose keys are held, assignment/deletion of parties who may be provided the keys, product/system cryptographic key management, actions on the audit log, audit profile management, and changes to the system configuration.
  - The product/system shall allow only specifically authorized subscribers to assume only those security administrative roles for which they have been authorized.
  - The product/system shall define a set of security administrative roles that minimally includes security administrator, system operator, cryptographic officer, and audit administrator.
- Non-Repudiation
  - The KRA shall be able to generate evidence of receipt for received transactions.
  - The KRA shall be able to generate evidence of receipt of registration or deposit of key recovery information from subscribers.
  - The KRA shall be able to generate evidence of receipt of requests from requestor.
  - The product/system shall generate evidence of origin for transmitted key recovery requests or responses.
  - The product/system shall provide a capability to verify the evidence of origin of information to the recipient.
  - The product/system shall provide a capability to verify the evidence of receipt of proof of receipt to the originator of message, i.e., recipient of proof of receipt.
  - The product/system shall provide the originator the ability to request evidence of receipt on information.

## Availability

The KRA shall provide a secure replication of any key recovery information stored.

## Protection of Trusted Security Functions

- The product/system shall provide a communication path between itself and local human subscribers that are logically distinct from other communication paths and provides assured identification of its endpoints.
- The local human subscribers shall have the ability to initiate communication via the trusted path.
- The product/system shall require the use of the trusted path for initial subscriber authentication.
- The product/system shall provide the authorized administrator with the capability to demonstrate the correct operation of the security-relevant functions provided by the underlying abstract machine.
- The product/system shall preserve a secure state when abstract machine tests fail.

## Policy

The KRA shall have a written policy based on the KRA Policy Framework. It shall operate in accordance with this policy.

## Registration Agent

Agents should protect all sensitive information from modification.

- Non-Repudiation – The RA shall be able to generate evidence of receipt for received transactions.
- Integrity – The RA shall be able to provide proof that information maintained has not been altered.

## Licensing Agent

Licensing Agents shall perform compliance audit of the KRA to ensure that the KRA operates in accordance with the KRA's stated policy.

## 8.1.5.12 Administration

Having good policies and technical solutions will not ensure the secure operation of a KMI/PKI or the validity of the subscriber certificates. An extensive set of operational policies and practices supporting the technical solutions also has to be in place. Historically, many problems found with infrastructure have not been with the technology, but with poor procedures, the operator did not know what to do in a given situation or the operator did not follow the proper procedures.

Administration of the infrastructure involves much more than the procedures to identify subscribers and create their certificates. It also requires managing the people, the components and the networks making up the KMI/PKI. Because of the wide range of activities that impact the KMI/PKI's security, the administration function is spread across a large number of people.

## UNCLASSIFIED

Key Management Infrastructure/Public Key Infrastructure (KMI/PKI)  
IATF Release 3.0—September 2000

Each of them must do their jobs correctly to have the level of trust defined in the policy. Specific tasks include:

- Policy enforcement, e.g., compliance audits.
- Administrating the network and system elements.
- Managing the technical security mechanisms for the infrastructure elements, e.g., administrating the computer's access control list, reviewing the audit files.
- Key and certificate accounting.
- Managing the cross-certification process.
- Managing the compromise recovery process.
- Defining and documenting operational and security procedures.
- Operator training.
- Physical and personnel security.
- Disaster recovery mechanisms.
- Maintaining availability.
- Managing the key recovery process.

Establishing the KMI/PKI's trust relationship via a set of policies (cf. Section 8.1.5.1, Policy Creation and Management) is the first step, but the trust model has to be continuously managed or it will become meaningless. The policies have to be translated into operational and security procedures for the specific technical solution employed within the KMI/PKI. These procedures provide a framework for the operators to administer the system. The procedures should cover all the normal processes in running the KMI/PKI and known exception and emergency activities. These procedures have to be documented and distributed to all appropriate KMI/PKI elements. The infrastructure should periodically re-examine and update the procedures as the policy changes, new processes are added, new exception cases are identified, new technical solutions are employed, or better ways of administrating the policy are found.

Once the infrastructure identifies and documents the procedures, the operators must be trained in the system policy and related procedures. Beyond the technical procedures necessary for their jobs, the operators must have an understanding of their responsibilities and limitations, and the security implications of not following the procedures. This process is open-ended because as the policy and procedures change, the operators need to be retrained.

The infrastructure has a responsibility to its subscribers and other domains to uphold its end of the trust relationship. This requires a mechanism to monitor the actions of every element in the KMI/PKI to ensure that they correctly implement the policy and procedures. Compliance audits, based on traditional concept of key management audits, are one way of tracking the subordinate elements. The root (or designated agent) periodically reviews each element to check the degree of compliance with the policy and procedures. The audit should also test little used and contingency procedures in order to determine if the operators would respond correctly. The results should identify and help correct problems with elements not properly implementing the procedures. Results should be available to other people in the trust relationship, e.g., domains that are cross-certified.

One of the most important extensions of the trust relationship is the addition of outside domains through a cross-certification. In effect, this gives every subscriber in the outside domain the same trust characteristics as an original member of the domain. This requires that the new domain have an equivalent level of assurance as the original domain (and vice versa). The only way to determine if this equivalency exists is to examine the two policies and determine whether they provide equivalent degrees of assurance. Standardizing on the format for documenting policies helps in comparing the policies by allowing a straightforward comparison of parallel certificate policy elements. One caution is that it is almost impossible to determine if the other domain actually implements their policy correctly (no independent compliance audit). The domains are trusting the correct enforcement of the other's domain security policy. Each domain has to monitor the other domain's performance and revoke the cross-certification link at any sign that it does not implement its policy correctly.

Since trust is the fundamental characteristic of the KMI/PKI, the physical and personnel security is important. A system operator can do extensive damage to the system and subscribers throughout the domain who rely on the certificates they authorize. Consequently, the people who have authorized access to the system must be trusted to do their jobs honestly, while all unauthorized subscribers are prevented from accessing the KMI/PKI.

Personnel security consists of both the hiring of the operators and their continued supervision. The owners of the infrastructure should perform some level of investigation (as defined in the policy) on their prospective employees to gain confidence in their trustworthiness. Periodic reinvestigations are necessary to maintain that degree of trust. If these reinvestigations or other actions bring their trustworthiness into question, the operator should be temporarily removed from access to the system. If further investigation confirms the suspicion, the keys and certificates they created may need to be revoked.

Physical security provides for the isolation of the KMI/PKI elements from access by unauthorized people. Protection is required for both the physical elements and their relevant KMI/PKI information. The policy should define the level of protection required. Because of the different sensitivities of elements within the infrastructure, the protection may vary. For example, the root might be located in a no-lone, i.e., an area where two-person integrity is required, zone protected with a 24-hour guard while a low level CA might only need a lockable protective container.

The technical security requirements must also be managed. These include the computers and networks that are used to implement and transport the infrastructure and its products. While these do not provide subscriber services, they are susceptible to attacks. If corrupted, they can negate other security mechanisms in the system. The infrastructure needs the same set of services, e.g., computer security, network confidentiality, intrusion detection, as other applications, so many of the solutions defined in Chapter 5 are applicable to the KMI/PKI.

The system administrators for the network, firewalls, and computer systems have to ensure that the underlying equipment works and provides the necessary security. The system administration should be a unique role and not done by an operator. The network administrator is responsible for providing network security services, e.g., authentication, access control, availability, and protection from network attack, and setting up the firewall. The computer system administrator is responsible for providing computer security services, e.g., least privilege, review audit files, access control, and virus protection. They have to install the computer equipment, setup operator accounts, define operator access privileges, monitor operator activities,

install new software, and install security software and patches. Administrator actions should be part of the compliance audit.

The KMI/PKI has to maintain its continuity in the face of an emergency that destroys infrastructure elements or during the routine elimination of existing infrastructure elements. That requires advance planning for each of the elements and the definition of appropriate disaster recovery mechanisms. Operators need to be trained in the recovery procedures and they should be tested as part of the compliance audit. The disaster recovery plans should guarantee the availability of the following services and information:

- Ability for subscribers to access certificates and compromised information.
- Ability to generate and distribute compromise information.
- Ability for subscribers to verify existing certificates.
- Archived records.
- Key recovery information.
- Authenticated copies of the old system parameters, e.g., root public key.
- Ability to reconstitute KMI/PKI with existing elements by creating new root and adding new elements as appropriate.

### 8.1.5.13 Requirements

Requirements related to the operation of the KMI/PKI include the following.

- The KMI/PKI shall ensure that a key or certificate request comes from an authorized source.
- Prior to issuing a key or certificate, the infrastructure shall verify that all the information within the request is valid.
- The CA shall authenticate a subscriber requesting a certificate to ensure that the correct public key is bound to the proper identity.
- The CA shall notify a subscriber when it has generated a certificate for that subscriber.
- With the exception of special circumstances, e.g., revocation due to firing an employee, the CA shall notify a subscriber when it has revoked the subscriber's certificate.
- The KMI/PKI will notify all subscribers of a revocation of a symmetric key.
- The KMI/PKI shall provide timely key and certificate revocation information to its subscribers.
- CAs shall provide their public key and/or public key certificates to subscribers in a secure and authenticated manner.
- A CA shall protect the private key material that it uses to sign certificates.
- The CA shall only use its signing private key material to sign certificates.
- If the KMI/PKI generates either symmetric keys or asymmetric key material on behalf of a subscriber, e.g., traffic encryption key or key agreement key material, the infrastructure shall ensure that the material is generated securely and securely distributed to the subscriber.
- If the KMI/PKI stores subscriber private key material for recovery purposes, the infrastructure shall ensure that this information is protected in storage and is only revealed to the

subscriber or to an authorized authority. It shall also ensure that the recovery key material is securely distributed to the subscriber or authorized authority.

- The KMI/PKI shall define a policy for the domain and ensure that all elements operate within the scope of that policy.
- The KMI/PKI shall account for the life cycle (ordering, generation, distribution, rekey, destruction and archive) of both symmetric key and asymmetric key materials and certificates.
- Proper technical and procedural controls shall be implemented to protect the components of the KMI/PKI.

## 8.1.5.14 Attacks and Countermeasures

### Attacks

Attacks that can be mounted against the KMI/PKI as a whole or to individual KMI/PKI components include:

- *Sabotage* – The KMI/PKI components or hardware token on which the subscribers or infrastructure elements keys and certificates are stored may be subjected to a number of sabotage attacks including vandalism, theft, hardware modification and insertion of malicious code. Most of these attacks are designed to cause denial of service. However, attacks such as hardware modification and insertion of malicious code may be used to obtain copies of subscriber or CA key material as they are generated, obtain information entered by the subscribers or operator such as a PIN, or cause known keys to be generated.
- *Communications Disruption/Modification* – Communications between the subscribers and the KMI/PKI components could be disrupted by an attacker. The disruption could cause denial of service, but may also be used by the attacker to mount additional attacks such as the impersonation of a subscriber or the insertion of bogus information, such as a key order, into the system.
- *Design and Implementation Flaws* – Flaws in the software or hardware upon which the subscriber depends to generate and/or store key material and certificates can result in the malfunction of the software or hardware. These malfunctions may deny services. The flaws may accidentally or be intentionally exploited to disclose or modify keys or certificates. Improper installation of the software or hardware may also result in similar consequences.
- *Operator Error* – Improper use of the KMI/PKI software or hardware by the operators may result in denial of service, the disclosure or modification of subscriber's keys and certificates.
- *Operator Impersonation* – It is possible that an attacker may impersonate a legitimate KMI/PKI operator. As an operator, the attacker would be able to do anything a legitimate operator could do such as generate key, issue certificates, revoke certificates and modify other infrastructure data.
- *Corruption or Coercion of the KMI/PKI Operator* – It is also possible that a KMI/PKI operator may be corrupted or coerced by an attacker to generate unauthorized key, issue certificates to an unauthorized subscriber, revoke certificates of legitimate subscribers and modify other infrastructure data.

## Countermeasures

Countermeasures that may be implemented to protect the KMI/PKI and its components from the attacks outlined above include:

- *Physical Protection* – Physical protection of KMI/PKI component hardware, communications link with other infrastructure elements, and/or hardware tokens will counter many of the sabotage and communications disruption related attacks.
- *Good Design Practices* – Concerns over flaws in the software and/or hardware design may be alleviated if good design practices are followed during the development of the software and/or hardware used in conjunction with the KMI/PKI.
- *Testing* – Testing of the software and/or hardware may also be used to counter attacks to the system that result from the exploitation of flaws in the system.
- *Training* – Training of the KMI/PKI operators and administrators is vital to eliminating or at least reducing the possibility of inadvertent attacks due to subscriber error.
- *Strong Authentication* – Strong authentication of the subscriber by the KMI/PKI components greatly reduces the possibility of impersonation attacks.
- *Access Controls* – Software or hardware based access controls may be implemented at the KMI/PKI components to limit the possibility that an unauthorized attacker will gain access to the infrastructure software or hardware.
- *Encryption* – Encryption of the link between the subscriber and the KMI/PKI components reduces the possibility that an attacker may eavesdrop on the communications and try to disrupt or modify the communications.
- *Contingency Planning/System Backup* – Backup of a subscriber's keys, certificates, and relevant software and hardware is the best mechanism for protecting against design flaws that result in system failure.
- *N-Person Controls* – Requiring multi-person control on sensitive PKI functions such as the process of bringing a CA to an operational mode, the generation of CA key material, etc. can limit coercion related attacks.
- *Auditing* – Auditing may not prevent attack, but it may be used to detect an attack and to identify the culprit. The presence of good auditing capabilities may also act as a deterrent to some attackers.
- *Personnel Selection and Screening* – Personnel chosen to perform KMI/PKI functions should be selected on the basis of loyalty and trustworthiness. People performing such functions should be adequately paid, and screened for a prior history which would indicate a pattern of untrustworthiness.

### 8.1.6 KMI/PKI Assurance

Section 8.1.1, KMI/PKI Introduction, addressed the KMI/PKI as a menu with a set of independent processes with independent solutions. However, a KMI/PKI's security is actually based on the *interaction between* all the processes. Since the intelligent attacker will always attempt the easiest attack that meets their goals, it makes little sense to have processes at vastly different levels of security. The effect is only to drive up the development and operational costs



without increasing the security posture. A better approach would be to determine the level of security needed for each application supported by the infrastructure and pick a set of solutions that correspond to that level of security.

Providing a high assurance KMI/PKI can be very expensive in both people and money. Cost effectiveness of many applications, like informal messaging, web browsing, or those handling low amounts of money, are very sensitive to PKI costs. For these applications, the KMI/PKI cannot cost more than a fraction of the potential loss from a successful attack. These applications may be willing to settle for a KMI/PKI that provides low cost certificates, but does not have all of the procedural and technical protections in place against certain attacks. In effect, KMI/PKI security is a form of insurance and employs the same cost considerations. A \$1 certificate is acceptable for protecting a \$100 transaction, but a \$50 certificate is not appropriate to protect the same \$100 transaction. Other applications may be willing to pay the added cost for better procedural and technical protections because the certificate is protecting more valuable information. A \$50 certificate might be acceptable if it is protecting a \$100,000 transaction.

There is much ongoing work in the standards community and the government in grouping the individual process solutions into fully developed architectures with common security standards. Among the groups working to define these standards are the IETF, FPKI, DoD, Canadian government, and commercial certificate providers.

## 8.1.7 KMI/PKI Solutions

Examples of KMI/PKI usage will be described to illustrate the practical aspects of system design. Three categories of systems—DoD, Government, and Corporate—each with important design and functional characteristics will be presented. Each of the examples of KMIs is actively involved in upgrading their Information Assurance (IA) assets and applications. The first category presented begins with summaries of the DoD Class 3 PKI and the FORTEZZA Public Key Infrastructure followed by a detailed description of the Target DoD KMI/PKI system showing its architectural development concerns, considerations and issues. The description is lengthy typifying the broad aspects of planning and considerations associated with a secure infrastructure implementation plus the added protections needed for processing classified information. This example demonstrates the challenge of designing a large system in today's environment. The DoD anticipates continued growth in the demand for security support for classified applications and Class 3 and Class 4 PKI capabilities. The Government KMI/PKI Solutions category will be presented next to show the many similarities with the DoD KMI/PKI despite its emphasis on UNCLASSIFIED//FOR OFFICIAL USE ONLY (U//FOUO) information. An example from the U.S. Government will be presented. The Federal KMI/PKI description includes the concept of bridging trust paths between PKI communities. The Corporate Solutions category is filled with a myriad of Commercial-Off-The-Shelf (COTS) products and services. Several are presented followed by a summary sketch of a corporate system. A short description using Kerberos for KMI security is provided to show some of the work being done in academia.

## 8.1.7.1 DoD Class 3 Public Key Infrastructure

### PKI Name

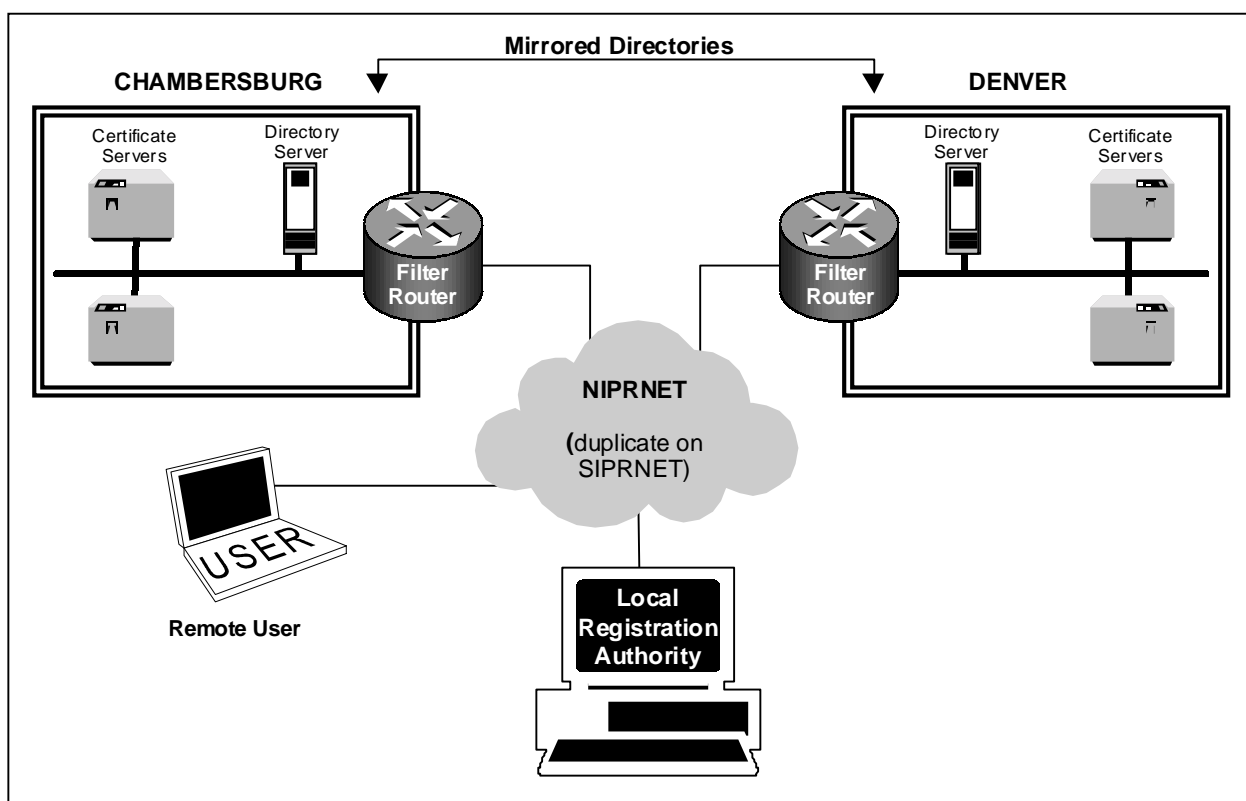
Department of Defense Class 3 Public Key Infrastructure (DoD Class 3 PKI). The original term, “medium assurance,” may be used interchangeably with the term, “Class 3.”

The following summary highlights the Department of Defense Class 3 Public Key Infrastructure (DoD Class 3 PKI) Solution, a forerunner of the DoD Target KMI/PKI described later.

### PKI Design and Operational Responsibility

The overall program management of all DoD efforts required to meet the goals and milestones in the DoD PKI Roadmap is the responsibility of the DoD PKI Program Management Office (PMO). The National Security Agency (NSA) is the PMO Program Manager with the Defense Information Systems Agency (DISA) providing the Deputy Program Manager leadership.

NSA is responsible for defining the security architecture and security criteria for the DoD PKI. This includes criteria for the components as well as their operation. NSA (or an approved National Information Assurance Partnership (NIAP) vendor) will evaluate the security of products and services employed in the DoD PKI. The architecture of the current DoD Class 3 PKI, minus the Root CA, is illustrated in Figure 8.1-14. The Class 3 PKI Root CA is located at the NSA Central Facility.



**Figure 8.1-14 DoD Class 3 PKI Architecture**

## PKI Subscriber Community and Applicability

One element of the Defense-in-Depth strategy is the use of a common, integrated DoD PKI to enable security services at multiple levels of assurance. The DoD PKI provides a solid foundation for Information Assurance (IA) capabilities and general-purpose PKI services (e.g., issuance and management of certificates and revocation lists in support of digital signature and encryption services) to a broad range of applications, at levels of assurance consistent with operational imperatives.

Classes 3 and 4 have been defined to support the protection of non-classified mission critical, mission support, administrative, or format sensitive information on open networks (i.e., unencrypted networks). These PKI Classes can also be used on closed networks (i.e., encrypted system high networks such as Secret Internet Protocol Router Network (SIPRNET)) to provide additional protection such as subscriber authentication and data separation/communities of interest (COIs). More specifically, Class 3 certificates and applications are appropriate for many business transactions, where the monetary value of the transaction or the sensitive or unclassified information protection is moderately high. By contrast, the Class 4 PKI products/services will be used to protect sensitive or unclassified mission critical information in a high-risk environment such as the Unclassified Internet Protocol Router Network (NIPRNET). Also, the Class 5 PKI products/services (still in the planning stages) will be used

for the protection of classified information on open networks or in other environments where the risk is considered high.

## **PKI Products**

The DoD PKI uses COTS products to keep up with technology evolution and develops Government-Off-The-Shelf (GOTS) solutions when necessary. The newness of standards and products, however, may cause some interoperability problems between vendors' products. The DISA and NSA actively work with vendors and standards communities to develop standard specifications and implementations that improve interoperability. The DoD is committed to ensuring that DoD specifications are consistent with the emerging commercial and National Institute of Science and Technology (NIST) Federal standards to support DoD interoperability requirements.

## **PKI Future Plans and Schedule**

The majority of activity to date in the DoD PKI arena has focused on understanding the technology, the standards, operational policy and procedural issues, and establishing the role of PKI relative to the rest of the IA Defense in Depth model. The experiences gained from the two major DoD PKI initiatives; the development and deployment of an operational FORTEZZA PKI, in support of the Defense Message System (DMS) and other FORTEZZA-enabled applications, and the pilot medium assurance PKI; have been instrumental in the development of the target DoD PKI architecture.

The DoD PKI will initially support three levels of assurance, defined as Classes 3 and 4 (formerly Medium and High) for the protection of unclassified/sensitive information, and Class 5 (for the protection of classified information on unencrypted networks). The long-term goal is to provide a Class 4 certificate to everyone within the DoD and—where appropriate—Class 5 certificates via the target DoD PKI. Each assurance level has its own set of requirements for technical implementation and process controls, which becomes more rigorous as the level increases.

The target DoD PKI shall employ centralized certificate management and decentralized registration, and use common processes and components to minimize the investment and manpower to manage and operate the PKI. Also, the target DoD PKI shall support a broad range of commercially based, security enabled applications and shall provide secure interoperability with the DoD and its federal, allied and commercial partners while minimizing overhead and impact to operations.

The DoD PKI program continuously tracks new/evolving IETF standards to ensure the most viable commercial standards are fully leveraged to support maximum interoperability in the future.

In addition, to ensure secure interoperability between DoD and its vendors and contractors, External Certificate Authorities (ECAs) will be established through a process that ensures the required level of assurance to meet business and legal requirements. ECAs will be approved by the DoD Chief Information Officer (CIO), in coordination with the DoD Comptroller and the Office of the Secretary of Defense (OSD) General Counsel.

The DoD PKI will be implemented in a series of steps to reach the final goals. The steps are highlighted .

- All DoD organizations must deploy registration applications for supporting the Class 3 (formerly Medium Assurance) PKI and the Class 4 (FORTEZZA-based) PKI.
- Protection of Category 1 mission critical systems on unencrypted networks using Class 4 certificates and tokens.
- Protection of Category 2/3 mission critical systems operating on unencrypted networks must use Class 3 certificates.
- Protection of Category 2/3 mission critical systems operating on unencrypted networks must use Class 4 certificates and tokens.
- Server Authentication.
- Client Identification (ID) And Authentication.
- Private DoD web servers access control software for Class 3 certificates.
- Email applications to facilitate digital signature processing of all individual messaging within DoD using Class 3 certificates.
- ID card processing software, building/facility access software, and workstation access software applications shall begin implementation for Class 4 certificates.

## Additional Information:

The following sources have more information on both the DoD Class 3 PKI products/services and the DoD PKI.

- Requesting Use of the DoD Pilot Medium Assurance Component of the DoD PKI (explains information/feedback to be provided to use the DoD Medium Assurance Pilot)
- DoD Medium Assurance Public Key Infrastructure (PKI) Home Page: <http://ds-2-ent.den.disa.mil/>
- US DoD X.509 Certificate Policy, version 5.0,13 December 1999; and DoD PKI Roadmap, Version3.0, 29 October 1999

### 8.1.7.2 FORTEZZA<sup>®</sup> Public Key Infrastructure

The following summary highlights the FORTEZZA<sup>®</sup> PKI Solution, a solution being used by the DMS.

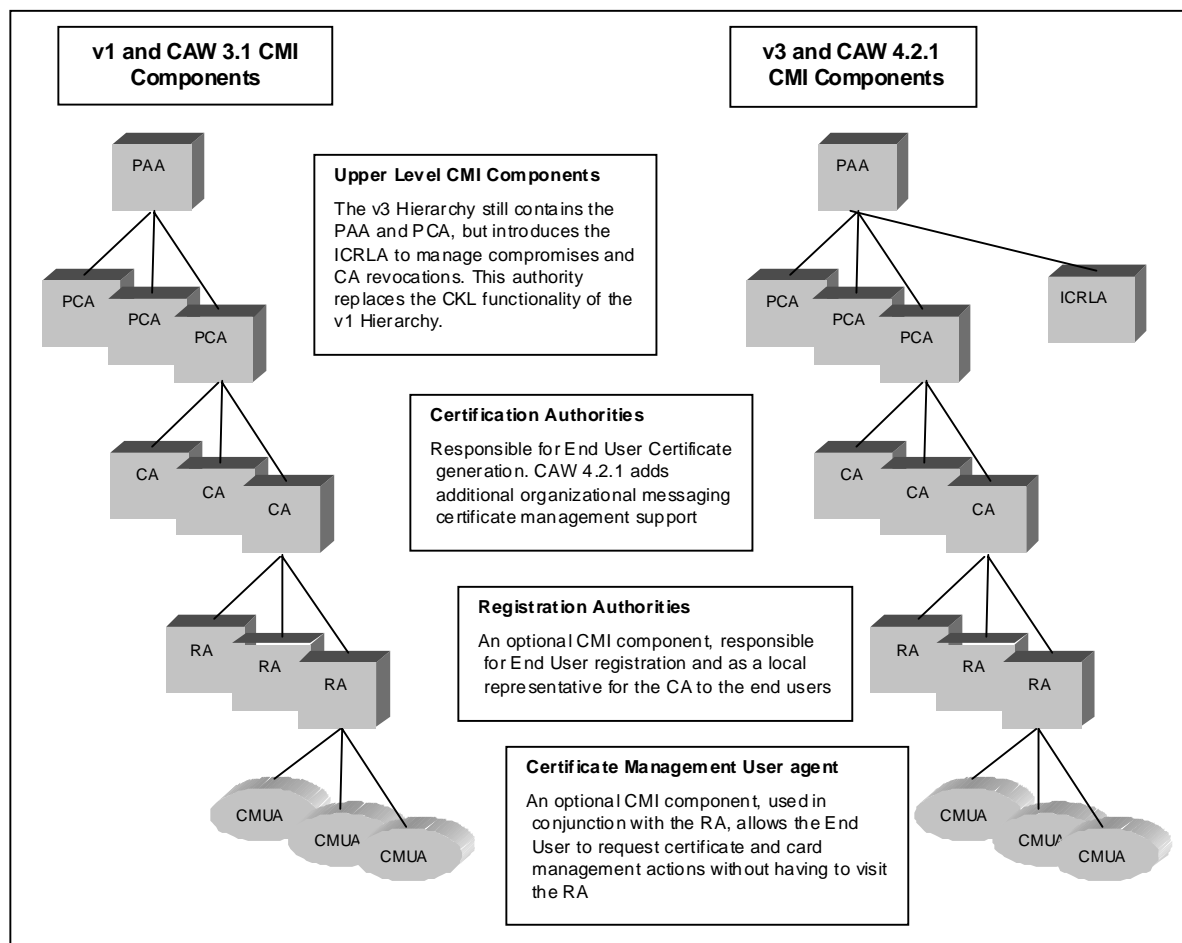
#### PKI Name

FORTEZZA<sup>®</sup> Certificate Management Infrastructure (CMI): A CMI differs from a PKI because it only includes the CMI and the policy associated with the CMI, not the directories where the public data items are posted.

## PKI Design and Operational Responsibility

The KMI Services and Workstation Technology division (NSA) is the Certification Authority Workstation (CAW) program management office responsible for its design, development, and testing. The Requirements and System Engineering division (NSA) and the Life Cycle Engineering and Standards division (NSA) are responsible for CAW life cycle support issues such as training, installation, upgrades, and maintenance; and the Electronic Key Management System Operations division (NSA) is responsible for the FORTEZZA CMI operations. Actual CAW training is accomplished via a combination of classroom, computer-based, hands-on, and on-the-job training, per policy, with the classroom training conducted by General Dynamics (CAW 3.1), Motorola (CAW 4.2.1) and Service Schools (both CAW 3.1 and 4.2.1).

Figure 8.1-15 illustrates the Policy Approving Authority (PAA), the Policy Creation Authority (PCA), the Indirect Certificate Revocation List Authority (ICRLA), the Certification Authority (CA), the Registration Authority (RA), and the Certificate Management User Agent (CMUA). Other CMI roles, not requiring dedicated workstations, are the System Administrator (SA) and the Information System Security Officer (ISSO).



**Figure 8.1-15 FORTEZZA CMI Components**

## PKI Subscriber Community and Applicability

The FORTEZZA PKI was targeted and is currently established to address certificate and security requirements of the DoD community, but its design and capabilities are also flexible to support civilian and commercial subscribers.

For DoD subscribers, the FORTEZZA PKI operates in compliance with Class 4 assurance policy due to its software design/development compliance with Trusted Software Design Methodology (TSDM) guidelines, its operation on a trusted operating system designed for the B1 level, implementation of high grade cryptographic algorithms and keys, and its strict use of hardware tokens for system infrastructure components. The certificates created and managed by the FORTEZZA PKI, when teamed with compatible applications, enable subscribers to apply all of the security services—authentication and identification, confidentiality, privacy or data integrity, non-repudiation, and access control—to unclassified as well as classified data. In addition, because of the high grade cryptographic algorithms, keys, and tokens that the FORTEZZA PKI implements, it is possible for applications to provide protection (authentication and confidentiality) for information to cross classification boundaries when such a crossing is already permitted under a system security policy (e.g. sending unclassified information through a High Assurance Guard (HAG) from SIPRNET to NIPRNET).

DoD organizations and customers of the FORTEZZA PKI can operate CAs in their local, decentralized environment and are responsible for complying with either NAG-69C, Information Systems Security Policy and Certification Practice Statement for Certification Authorities, (for X.509 v1 use with CAW 3. 1, CAW 4.2. 1) or DoD Certificate Policy, Version 5.0 (for X.509 v3 use with CAW 4.2.1).

## PKI Products

The FORTEZZA PKI supports secure DoD transactions across existing national and global information networks (e.g., Internet) and allows them to be protected from threats from other subscribers of the global information network. The functionality of the current FORTEZZA PKI, based on CAW 3.1, supports:

- UNCLASSIFIED//FOR OFFICIAL USE ONLY (U//FOUO) and Classified environments on the same CA platform.
- Performs trusted downgrade of information between different classification levels of network(s)/account(s).
- Creates and manages X.509 v1 certificates.
- Creates and manages v1 CRL.
- Creates and manages card, certificate, and DN in a flat file database.
- Manually posts certificates, CRLs, Compromised Key Lists (CKLs) to X.500 Directory System Agent (DSA)
- Processes MISSI Management Protocol (MMP) messages from other networked devices.
- Implements Message Security Protocol (MSP) 3.0.
- Manages backup data for certificates, CRLs, and CKLs.

## UNCLASSIFIED

Key Management Infrastructure/Public Key Infrastructure (KMI/PKI)  
IATF Release 3.0—September 2000

The CAW 3.1 can be configured to serve as a Policy Approving Authority (PAA), Policy Creation Authority (PCA), or CA.

The optional Registration Authority using the Motorola Registrar product, provides a cost effective alternative to dedicated CAWs for both multiple subscriber registration and routine subscriber certificate update tasks. Registrar 4.2 is available now to support CAW 3.1 as well as CAW 4.2.1 when it is fielded.

The optional Motorola Certificate Management User Agent (CMUA) resides on subscriber Windows NT platforms and further off loads subscriber registration and maintenance functions from the Registrar. This product is available now to support CAW 3.1 as well as CAW 4.2.1 when it is fielded.

## PKI Future Plans and Schedule

The FORTEZZA PKI (X.509 v1 certificates only) has been operational since March 1995. The current PKI (operational since January 1998) is based on CAW 3.1. An upgrade to CAW 4.2.1 began in March 2000 for the PAA and PCAs and staggered upgrades of the CAs in the field. The March 2000 upgrade is backward compatible to CAW 3.1 functionality and its X.509 v1 certificates. The March 2000 upgrade also provides a totally new software design and code based on TSDM Level 3 guidelines, a new and improved Graphical User Interface (GUI), a relational database, automatic posting of information to a public directory, management of multiple hardware and software tokens, programmable X.509 certificate extensions for flexible security policies, X.509 v3 certificates, v2 CRLS, and Indirect Certificate Revocation Lists (ICRLs).

Plans are underway to develop and field a future CAW version to provide support for Software FORTEZZA technology and capabilities.

## Additional Information

Additional information can be found in the following documents:

- Interim Operational Security Doctrine for the Sensitive-But-Unclassified (SBU) FORTEZZA Card, 18 February 1998.
- Interim Operational Security Doctrine for the FORTEZZA For Classified (FFC) FORTEZZA Card, June 1998.
- NAG69B, Information Systems Security Policy and Certification Practice Statement for Certification Authorities, 24 October 1997 (for X.509 v1 with CAW 3.1 and 4.2.1).
- NAG69C (Replacement for NAG69B, Pending final approval at NSA)(for X.509 v1 with CAW 3.1 and 4.2.1).
- DoD Certificate Policy, Version 5.0, (for X.509 v3 with CAW 4.2.1).
- FORTEZZA Public Key Infrastructure (PKI) Concept of Operations (CONOP), Version 1.8, 7 January 2000.
- Certificate Management Infrastructure (CMI) Transition Plan, Version 2.0, 23 November 1999.

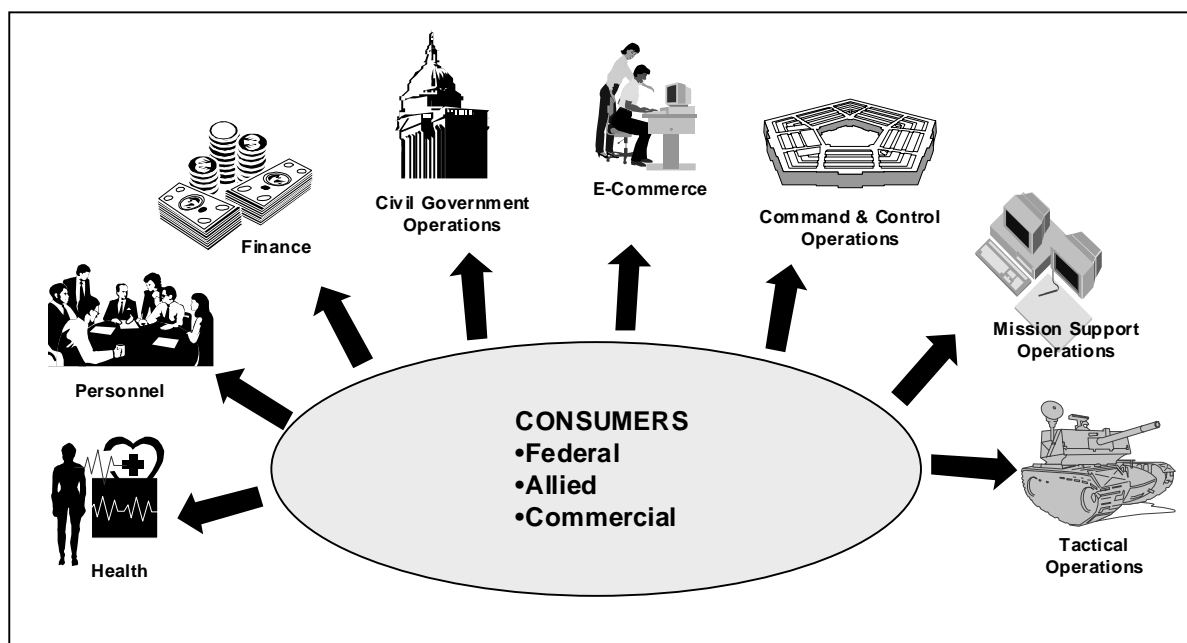


## 8.1.7.3 DoD Target Key Management Infrastructure

Throughout the following text, KMI is used interchangeably with KMI/PKI.

### 8.1.7.3.1 Background

The people, programs, and systems that carry out or support the broad range of missions of the DoD perform a variety of activities. These diverse activities, depicted in Figure 8.1-16, represent an ever-expanding need, and thus role for IA capabilities in DoD operations. Traditionally, DoD has addressed these needs with stand-alone cryptographic components. In today's Information Technology-rich environment, DoD's IA needs are being addressed with security features integrated into the many communications and information processing system components used by the DoD. These include workstations, guards, firewalls, routers, in-line network encryptors, software applications, and trusted database servers. The deployment of the large numbers of these security-enabled components (both traditional cryptographic devices and integrated IA features) is placing an increasing burden on the network infrastructure that provides KMI products and services.



**Figure 8.1-16 Operational Activities Supported by the KMI**

The DoD KMI is a foundational element for a secure IA posture in the Defense Information Infrastructure (DII) and the broader National Security community. The DoD is taking an aggressive approach in acquiring a KMI that meets the requirements for all IA key management needs. The DoD KMI program, supported by the Services and Agencies, Joint Staff, and the DoD contractor community, is addressing this critical need.

The state of the current key management systems creates compelling reasons for modernizing the DoD KMI.

## UNCLASSIFIED

Key Management Infrastructure/Public Key Infrastructure (KMI/PKI)  
IATF Release 3.0—September 2000

- *Infrastructure of Independent Stovepipes*—The current key management environment is made up of separate and independent infrastructures that provide and manage their own set of security products. These systems will become increasingly cumbersome and costly as new technology and their attendant security solutions continue to advance and the resources needed to operate them decline. This key management environment is compromised of several unique solutions built for specific product lines. While the solutions satisfy unique security needs, they each require different tools and training in order to obtain their respective products and services, imposing an unwarranted strain on resources.
- *Inefficient Expansion of New Capabilities*—Adding new key management capabilities has frequently required integration of new capabilities into existing systems that were not designed to perform the new functions, or the creation of new, independent systems to provide the needed support. One recent example is deployment of a totally separate (stovepipe) network infrastructure to support DoD's use of PKI-based security products. While this is an example of limitations of the existing KMI structure, other programs are running into the same issues. This is slowing down DoD's ability to respond to new requirements, and demanding more resources for supplying cryptographic key products to support its missions.
- *Common Functions and Operations*—Although created independently, the existing systems contain many common threads (e.g., registration, ordering, and distribution) that could logically be combined and offered as a unified set of processes. Not only has the key management community recognized this fact, so has the DoD Joint Staff. They have identified a unified KMI as a critical system infrastructure, needed to support key and certificate management approaches for mission critical, logistic, and administrative systems.
- *Opportunities for Applying New Technologies*—Several of the KMI systems have been in existence for a number of years and are in need of upgrade to take advantage of modern communication technology. This technology area has advanced significantly in recent years, providing the market place with many new and worthwhile, applicable techniques that would greatly improve efficiency and performance.

Given the critical importance of key management, applying modern technology within a sound IA systems approach is imperative. The KMI initiative focuses on unifying the disparate key management systems within a single, modern architecture—one that is modular, flexible and extensible—one that will eliminate redundant resources associated with operation, maintenance and training, resulting in substantial cost savings.

Commercial security technology using public key cryptography for U//FOUO requirements is rapidly becoming the largest "application class" that must be supported by the DoD KMI. However, requirements for support of classified applications are also projected to continue to grow significantly as new classified solutions such as secure wireless and Global Positioning System modernization are implemented. This creates the need for a more encompassing key management paradigm. It is the intent of the KMI to enhance the DoD's capability to support these mission-critical requirements. The DoD KMI Program will unify these many disparate key management systems within a single, modern framework, introduce additional key management capabilities to support the continued expansion of KMI services that are projected, and address the Congressional mandate to reduce operational costs associated with the KMI.

## KMI Products and Services

KMI, as described herein, refers to the framework and services that provide registration, enrollment, generation, production, distribution, control, and tracking of the broad range of KMI products needed by the DoD. A critical challenge for the KMI will be to provide continuing support for existing products and services, as well as emerging security solutions. At a minimum, the following product categories will be supported.

- Human-readable cryptographic products (e.g., code books, one-time pads, authenticators, key lists).
- Symmetric cryptographic key for point-to-point and net use and for use in wireless products.
- DoD Class 3 PKI Root CA.
- Asymmetric cryptographic products.
- Electronic certificates (e.g., signature, attribute, and key exchange) used in a multitude of applications to implement security functions such as identification and authentication (I&A), access control, integrity, confidentiality, and non-repudiation.
- Key management documentation (e.g., policy documents, equipment operator manuals, and specifications) needed in support of the cryptographic user community.

The Target KMI provides the framework and services that unify the secure creation, distribution, and management of these products. The DoD KMI will enable the provisioning of these services for military, intelligence, allied government, contractor, and business customers. A baseline set of key management services offered by the KMI to support the user community includes the following.

- *Registration*—Identifying, in an authenticated manner, individuals, or system entities (either internal or external to the KMI) and their related attributes.
- *Enrollment*—Authenticating the establishment, modification, and deletion of privileges for individuals, system entities, or organizations.
- *Ordering*—Requesting cryptographic product (e.g., keying material, certificates, manuals, etc.) to support a security application.
- *Generation*—Generating cryptographic products (e.g., symmetric key, asymmetric key and/or a public key certificate) by a security infrastructure element.
- *Distribution*—Providing physical and electronic products—to include rekey—to the user in a secure, authenticated manner.
- *Policy Management*—Managing and enforcing both policy and procedures for operating the KMI in a trusted and secure manner.
- *Trust Extension*—Reviewing and ruling on issues of cross certification or bridging with other key management infrastructures.
- *Archiving*—Providing for long time storage and retrieval of important data that may not be immediately accessible to on-line users of the system.
- *Accounting*—Tracking the location and status of cryptographic products.
- *Key Recovery*—Recovering encrypted information when the intended decryption key is unavailable.

## UNCLASSIFIED

Key Management Infrastructure/Public Key Infrastructure (KMI/PKI)  
IATF Release 3.0—September 2000

- *Compromise Management*—Providing notification of compromised keys and invalid certificates in a timely and authenticated manner.
- *Audit*—Supporting periodic security evaluation of KMI operations.
- *Library*—Providing access to key management reference documents and information.
- *Destruction*—Destruction of certificates and keying material.

## Planned Evolution

The DoD KMI will be implemented as a series of evolutionary phases culminating in a re-designed, unified architecture. Strategic and architecture planning will require in-depth coordination with KMI government and commercial partners. Every 18-24 months, a new Capability Increment (CI) will be delivered to operational users taking into account new and updated user operational, security, policy, and technology requirements, and programmatic opportunities. Timing of the capability increments is critical to ensure optimum synergy and cohesion with the individual systems in the DoD KMI architecture. For each CI, the Target KMI will be redefined to be consistent with current and projected operational/security needs and technology advances. The updated Target KMI definition will be used for programming and budget planning for the products and services needed to realize the Target KMI. This approach requires sustaining system engineering and development resources, and wide Service/Agency/Organization support for the acquisition, deployment, and operations of each CI.

The KMI utilizes a wide variety of existing networks and workstations to fulfill its mission and is being designed to implement as many KMI-wide functions as possible on COTs platforms. Initial deployments of the KMI will be structured as separate KMI functions for each security classification domain. However, as the system evolves it will transition to a structure that allows the transfer of appropriate data between domains. Using this approach, most KMI functions will operate on a single-level (commercial) System High platform at client manager nodes as well as in the centralized portions of the system infrastructure.

## Goals and Objectives

A number of goals have been identified for the KMI based on user community input security, advancing technology and the reality of a shrinking budget. These include the following.

- **Transparency**—While some functions within the KMI inherently require direct operator or user interaction, the KMI will automate as many operations as possible. KMI-aware devices will interact with the KMI, transparent to the user.<sup>1</sup> Current, manpower-intensive operations (including accounting and archiving) will be automated and transparent to KMI users.
- **Ease of Operation**—The Target KMI will provide simplified, intuitive, and consistent interfaces for users to obtain KMI support for the ever-increasing range of PK functions. Users will have standard Web browser access to the KMI—with screens tailored based on

---

<sup>1</sup>While the KMI can provide secure infrastructure capabilities to enable this transparency, modifications to KMI-aware devices are also required to add functionality that can realize this transparency.

their identity, role—and authorized capabilities, and KMI products and services tightly integrated into their mission planning and system management capabilities.<sup>2</sup>

- **Access to Needed Information**—The KMI will offer direct, on-line access on all relevant policy information and to operational information (e.g., inventories of keying materials and cryptographic devices) to ensure that policies are carried out appropriately. Customer Support will be provided 24 hours a day, 7 days a week to assist users with KMI-related issues.
- **Reduction in User Manpower Support Needs**—Continued proliferation of cryptographic devices (user terminals, network servers, security-enabled network devices) and projected wide-scale deployments of PKI-enabled software applications will continue to increase user manpower burdens to get KMI products. The Target KMI will reduce this burden with its greater use of commercial standards and products.
- **Responsive Policies and Doctrine**—Uniform, national level and DoD-wide policies, doctrine, practices, and procedures will be established in joint-community forums to ensure interoperability and consistency of Joint operations at the organizational level. They will be coordinated and issued prior to deployment of cryptographic equipment.
- **More Efficient Use of KMI Operator (internal) Support Needs**—Continued proliferation of cryptographic devices (user terminals, network servers, security-enabled network devices) and projected wide-scale deployments of PKI-enabled software applications will continue to increase demands on KMI operator manpower needed to generate and produce KMI products. The Target KMI will be more efficient than the existing KMI, allowing them to deliver products faster and respond more quickly to new requirements.
- **Enhanced Security**—Delivery of all orders will be available securely and directly to the end-user or end-user devices that require them. The KMI will be built on authentic, universally accepted identities for all users, operators, and devices. Standard tools and tool kits will be provided by the KMI to ensure that all KMI-relevant operations (e.g., key exchange, rekeying, and certificate path validation) are performed correctly.

## General Features of the Target KMI

There are several pervasive characteristics of the Target KMI. These include the following.

- **Modularity**—The Target KMI, while still being refined, is based on a modular structure that will enable adequate flexibility to ensure that it can evolve over time. It will immediately leverage existing key management system capabilities and commercial components (e.g., commercial certificate authority workstations, directory systems) in the baseline implementation and incrementally evolve the capability as commercial technology matures. The KMI capabilities will evolve, taking advantage of commercial technologies; a strategy that requires a DoD enterprise-wide standards approach, and a coordinated process within DoD to influence the direction of commercial standards bodies to incorporate features important to the DoD.

---

<sup>2</sup>In a similar manner, this goal can only be realized with enhancements to mission planning and system management components.

## UNCLASSIFIED

Key Management Infrastructure/Public Key Infrastructure (KMI/PKI)  
IATF Release 3.0—September 2000

- **Automated Service**—The KMI will offer a well-defined set of KMI products and services, with an established set of delivery mechanisms and interface standards for “last-mile delivery devices,” clearly defining how KMI products will be delivered.
- **Key Delivered Directly to End-Devices**—The KMI will evolve toward electronic delivery of key, with delivery directly to end-devices. The KMI will provide tool kits that can be used to KMI-enable devices and operational support systems to take full advantage of the advanced features and capabilities that the KMI will offer.
- **Common Management Functions**—The KMI will introduce a set of common management functions that will enable consistent KMI operations currently provided by the various existing stovepipe KMI systems. It will augment these with a set of primary services (e.g., registration, common ordering, and key recovery) that will enable common KMI interactions for users and KMI-aware devices to obtain the specific KMI products or services they require. It will also incorporate both functional and physical modularity to facilitate an orderly introduction and enhancement of operational capabilities throughout the KMI’s life cycle.
- **On-Line Customer Support and Library Access**—The KMI will include an on-line repository to provide authorized KMI users and managers with a complete catalog of KMI products and services, test results of commercial IA products, electronic versions of current policies, manuals, advisories, and inventory status for deployed KMI-relevant devices and KMI products (including those of Allies and Coalition partners).
- **Leveraging Existing KMI System Investments**—The KMI encompasses products and services provided by the Electronic Key Management System (EKMS) physical key management capabilities and operational public key infrastructure capabilities. These provide a wide range of cryptographic keys for traditional symmetric key systems and key pairs and certificates for public key systems. The Target KMI provides the framework and services that will allow DoD to incorporate the existing KMI systems into the Target, thus improving the existing underlying system infrastructure that provides security services to military, intelligence, allied government, contractor, and business customers.
- **National Level Policies**—There are many KMI challenges facing DoD, and it is anticipated that the implementation of DoD KMI will result in changes to areas such as national cryptographic policy to

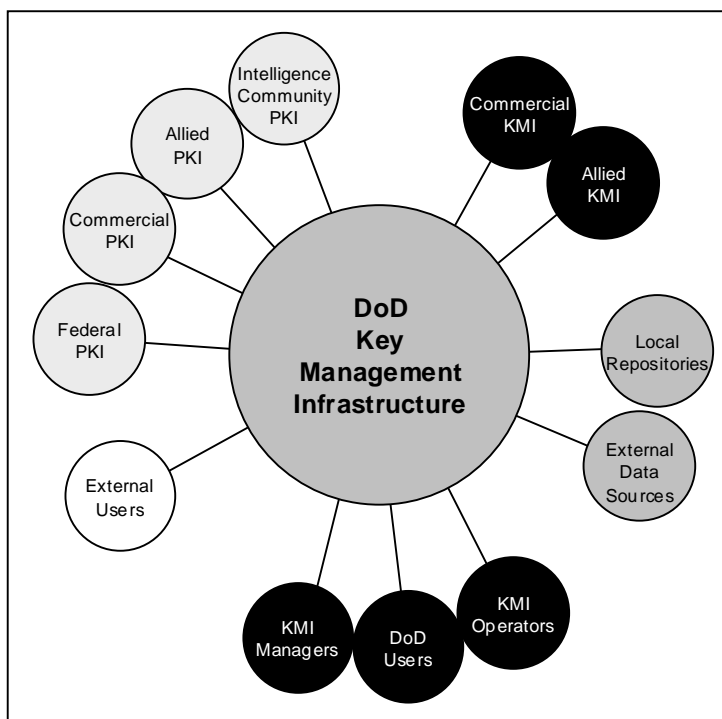


Figure 8.1-17 DoD KMI System Context

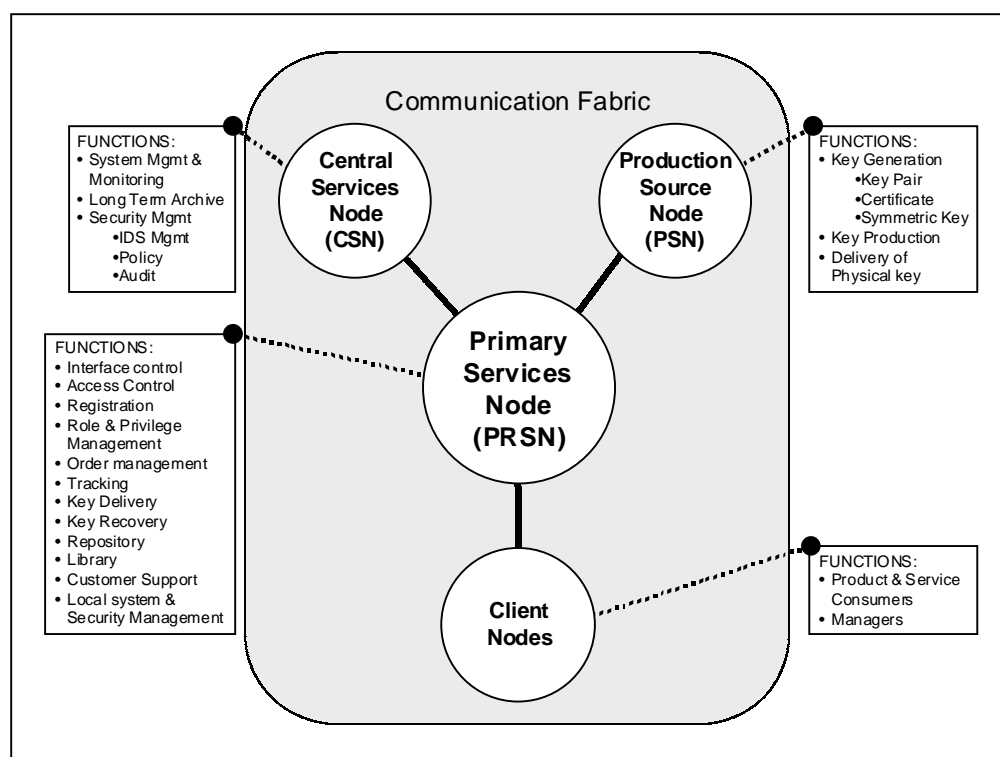
better coordinate the handling of classified and non-classified key management data.

## System Context

The KMI interacts with a number of external components and systems to perform its intended functions. Figure 8.1-17 provides a context diagram for the KMI system capability. One of the primary capabilities is to interact with the users it is intended to serve. It must also interact with external Federal and commercial KMIs and PKIs. It interfaces to external data sources, including local user community repositories and external data sources such as the Defense Eligibility and Enrollment Reporting System (DEERS) database that contains personnel information that may be accessed during registration of some end users.

### 8.1.7.3.2 DoD KMI System Context KMI Nodal Architecture

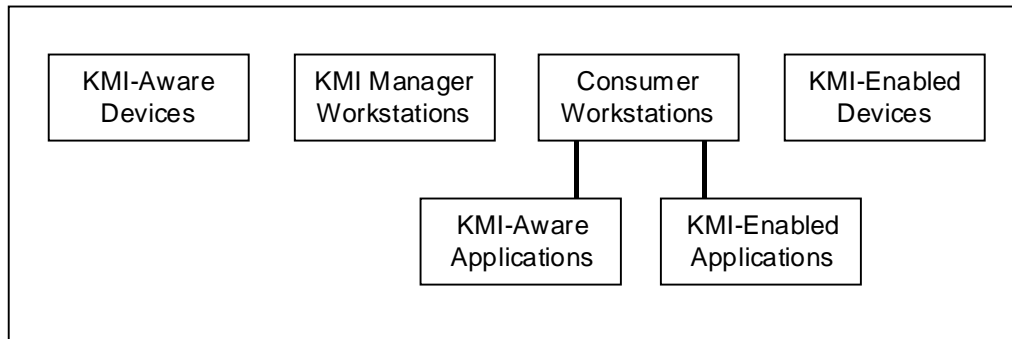
The Target KMI architecture consists of four types of functional nodes, as depicted in Figure 8.1-18. Their interconnectivity and summary of the major functions of each of these nodes is included in the figure, and discussed in more detail below. Section 8.1.7.7 identifies the major documents that describe the Target KMI in detail.



**Figure 8.1-18 Nodal View of the Target KMI**

## Client Nodes

The Client Nodes represent both the consumers of KMI/PKI products and services and the workstations that support the various KMI/PKI managers. Figure 8.1-19 provides a breakdown of several generic types of clients. Client nodes, also referred to as *end entities*, include stand-alone cryptographic devices, devices that incorporate security features that rely on key management services (e.g., security features within a router), and workstations that use software applications that require KMI support.



**Figure 8.1-19 Breakdown of Client Nodes**

## Primary Services Node

KMI users—whether they are humans, devices, or applications—obtain their products and services from a Primary Services Node (PRSN). The PRSN provides common management functions in a server-based architecture and provides its required services in multiple classification domains. The PRSN provides to the client node components, unified and transparent access to all of the different production sources and delivery of KMI products and services to consuming applications, directly or through an intermediary. As implied in the figure above, the PRSN is also the node that handles user access. When KMI products are requested, the PRSN will forward the request to the appropriate Production Source Node (PSN) for generation and production. If the product can be delivered electronically, the PRSN will forward it on to the Client node.

## Production Source Node

The Production Source Nodes (PSNs) are responsible for the generation and production of KMI products. These products will be created at the request of PRSNs. If a physical product is needed, the PSN is responsible for delivering the product directly to the client node. PSNs are separated from the common management functions of the PRSN, but interface via available communications networks to the management infrastructure provided by the PRSN. The Electronic Key Management System (EKMS) Central Facility and Key Processor (KP), the existing physical systems, and the PKI CA are examples of current KMI systems that provide functionality associated with a PSN. The Target KMI architecture has adopted a modular structure specifically to accommodate the modification of existing, or addition of new production sources.



## Central Services Node

The Central Services Node (CSN) provides overall system management and monitoring functions for the system infrastructure. In the Target KMI, the CSN will provide the long-term system archive as well as the master KMI database, and will replicate data to the individual security enclaves of the PRSNs. The CSN will also handle overall system infrastructure security management, including IDS oversight, audit data collection and analysis, long-term archiving, policy management, and system health monitoring.

## General Deployment Considerations

The Target KMI will be deployed as modular sites consistent with the nodal architecture discussed above. There will be one CSN and a physically isolated hot backup to mitigate risks of natural disasters interrupting operation. There will be several PRSN sites in strategic locations across CONUS.

Each will be capable of serving as a back up capability to other PRSNs, with automated cutover capabilities available to ensure uninterrupted service to KMI clients. Deployable versions of PRSNs will be established in sites outside CONUS to minimize network connectivity issues for operations in various theaters. Typically, these sites will reach back to the CSN and PSN located in CONUS. To the extent that PSN capabilities are needed to support these deployed sites, a Black PSN will be available to provide the capability (using stored materials that can be transferred via physically and/or electronically-protected means), minimizing the risks of operating in potentially hostile environments. The deployed PRSNs will also include basic CSN provisions to facilitate operations when connectivity back to CONUS is impaired or unavailable.

The KMI will utilize the communication channels already serving its customers in other capacities. The KMI will rely on existing communications paths for connectivity within the system. The KMI will also support dial-in capability through secure terminal equipment. Once connections are established, the interfaces and functionality will be the same as that available when connecting to the KMI through a data network.

### 8.1.7.3.3 Perspectives on KMI Operations — An External KMI Perspective

This section provides an operational overview of major Target KMI functions from the perspective of users and managers of KMI products and services. Further detailed descriptions of these operations can be found in the Target KMI Concept of Operations Document.

The Target KMI is designed to automate operations to the extent that it is feasible and prudent. For those operations requiring human intervention, the KMI provides standard operating procedures for a range of user and manager functions (referred to as common management functions). In general, KMI user and manager operations will be performed as local client workstations interacting with server capabilities in the PRSN. In general, a KMI user or manager will insert their KMI token into their workstation, log into the PRSN, request a

particular KMI function and be connected to the appropriate server.<sup>3</sup> Where feasible, the PRSN will provide intuitive screens with pull down menus tailored to the specific role(s) and privileges of the requester.

## Registration

Registration is the process that allows an end entity to become *known* to the KMI. It establishes the identity of the end entity that the KMI asserts for all of its operations. Registration also results in the generation of an identity certificate and the creation of a token that is delivered to, and remains in the possession of, the registrant. In general, KMI registration is a decentralized process that is performed by a number of Registration Managers (RMs), including RAs and LRAs. Within the context of the Target KMI architecture, the RM is a client node manager, and is typically someone who is located close to the user.

The DoD PKI Certificate Policy (CP) establishes the requirements and policies that are used during registration. In a typical scenario, a registrant appears in person before a RM and presents credentials of their identity as required by the appropriate CP and CPS. To register devices, the device sponsor or component administrator submits appropriate documentation about the device to their LRA. The RM logs into the PRSN using a KMI token to establish privileges and accesses the Registration Server. The RM validates that the information provided by the individual agrees with independent identity data obtained from an independent external data repository (e.g., DEERS database or a repository provided by their department, agency, or organization).

## Enrollment

Enrollment is the association of privileges with an individual's KMI identity by a KMI Privilege Manager (PM). Enrollment enables KMI users and managers to conduct transactions for which they have been granted privileges. Each KMI operator and each Client Node manager has a defined role (or set of roles) in the KMI, and roles determine the scope of privileges within the security infrastructure. For example, the role of a RM, like an LRA is to register users in the system. Other managers, such as User Representatives or Product Requesters may, order keys, certificates, or other services from the KMI on behalf of registered users. A PM performs the function of defining roles, allocating privileges to those roles, and assigning roles to individual managers.

## Request and Tracking

The process of requesting KMI products and services and then tracking the status of those requests is structured in a fashion similar to registration and enrollment. Provisions are also included for direct requests to be made from KMI-aware devices that have been configured to perform KMI transactions transparent to users and operators. An authorized KMI end entity inserts their KMI token into their workstation and accesses the PRSN Common Ordering Manager. They may choose to access a catalog of all KMI product and services offerings on-line in the KMI Library. They are also offered a menu of templates for each KMI service and

---

<sup>3</sup>While reference is made to specific "servers," in actuality it represents functional capabilities of the KMI node referenced.

product that they have been assigned a privilege. The templates are tailored to limit selection to only those options that they have been granted privileges. They can either retrieve an existing request through the template and modify the data for resubmission or access a blank template. Once the request is completed, they submit it to the PRSN.

Tracking of orders is performed in a similar fashion. Each order is given a tracking number that can be referenced. An authorized operator can access a list of all pending orders. They can choose to query for status, update, or cancel a request. They can also choose to remain on-line while the status is requested, or select to have the PRSN send a notification of the action when it is available.

## Distribution

This process arranges for the transfer of KMI products from the KMI to end users or intermediaries in a secure and authenticated manner. There are two basic types of KMI products that are distributed. The first includes physical products (e.g., hard copy codebooks, canisters of hard copy key materials, and tokens). These are distributed through protected shipping channels (ala the Defense Courier System). A goal of the Target KMI is to reduce to the extent operationally acceptable the amount of these materials. The preferred means of distribution is protected electronic delivery. When a KMI product is available for distribution, it can be “pushed” automatically to the intended recipient. The PRSN includes an electronic vault for intermediate storage of Black KMI products that have been generated previously. The KMI provides the capability for authorized users to “pull” materials from the vault. The vault also serves as a rapid access source for products that the KMI will deliver (or “push”) to end entities.

## Key Recovery

Key recovery capabilities allow a means for authorized KMI users to access KMI products associated with an encryption process (referred to as Key Recovery Information or (KRI)) in the event that key is lost or otherwise unavailable. There are two general applications for key recovery. One is to enable local information owners to access information that is protected when a key is lost. The other is a central capability to provide KRI to other authorized individuals based on national policies for key recovery.

## Revocation

Revocation is used in normal operations as individual responsibilities and privileges change, resulting in the need to invalidate individuals' KMI roles and privileges. It is also a critical component of recovery in the event that sensitive KMI materials of an individual, a KMI manager, or an internal KMI operation have been compromised or are suspected of being compromised. The process for requesting a revocation is performed in the same manner as KMI product and service ordering. An authorized KMI manager inserts their KMI token into their workstation and logs into the PRSN. Their workstation will access the Compromise Recovery Agent within the PRSN, that will validate the manager's identity, and the role and privileges associated with that identity. They are also offered a menu of intuitive templates to allow them to accomplish a revocation request. The templates are tailored to limit selection to only those options that they have been granted privileges. The KMI processes that request, and

activates mechanisms automatically to prevent any operations using the revoked KMI materials.

#### 8.1.7.3.4 System Operations—An Internal KMI Perspective

While the previous section highlighted critical KMI system operations from the perspective of KMI users and managers, this section provides an overview of the internal operations of the Target KMI to support system functions. The KMI is designed to provide a set of common management functions to provide a uniform, consistent, and intuitive interface to KMI users and managers.

KMI manager and end user workstations are structured as “light clients,” using commercial web technologies to support transactions with servers provided in the PRSN. This allows system enhancements to focus on updates to these servers, minimizing reconfiguration of RM software. From the perspective of the KMI internal operations, the KMI end entity uses a workstation and KMI token to access the PRSN. The connection is secured using the token as a basis for establishing identity and securing the transactions. The PRSN Access Manager validates the end entity’s identity, role(s), and privileges before access is granted to any other KMI resources. For all operations, each server within the KMI will verify the privileges for the identity represented in the token and whenever feasible will provide tailored screens with pull-down menus for the entity to select any authorized operation desired. Archiving of audit information for all interactions will be maintained automatically by the PRSN. Tools will be available to allow authorized users and managers to query the audit information.

### Registration

Registration by its nature requires involvement of users and operators. Registration allows an individual or device to receive a PKI identity. The RM accesses the PRSN and logs into the Registration Server. Using screen menus tailored for registration of the type of entity being registered, the RM enters the required identity information. The workstation, via the PRSN, accesses the external repository for information to be validated. It presents this information to the RM, annotating possible discrepancies. Once the RM accepts the identity as valid, the workstation develops an identity certificate.

There are several concepts still being considered for processes at this point. The scheme currently used is for the token to generate a public and private key pair. Other options are for the end user workstation or the CA to generate the pair. When the token generates the pair, the token transfers the public component to the RM workstation, that in turn forwards it along with a certificate request through a PRSN for registration to a PKI CA PSN<sup>4</sup>. The PRSN assigns a KMI Unique (KU) identifier to the identity. The CA creates/signs an identity certificate, updates the appropriate directory, and returns the certificate to the RM workstation. The RM workstation loads the certificate onto a token and the RM issues the token to the user. All tracking and audit information is performed automatically by the PRSN and CA PSN, as appropriate.

---

<sup>4</sup>In selected operations, the private key is also transferred in a secure fashion to the CA (via the PRSN) to support future key recovery operations. Private keys associated with identity certificates are NOT escrowed.

## Enrollment

Using a KMI token to establish identity, the PM accesses the PRSN and logs into the PRSN Enrollment Server. Enrollment allows an individual or device to receive encryption keys. The PM then inserts the token for the end entity being assigned KMI roles and privileges. The Enrollment Server provides menu screens for the PM to select the operations desired. This includes the update of role definitions, privilege assignments to roles, and identities assigned to roles. All of the PM interactions will be automated and updated into the KMI Library repository that stores enrollment status information.

## Request and Tracking

An authorized KMI user or manager can access the PRSN and log into the Common Ordering Manager to request KMI products and services and to obtain status of requests that are being processed. The Common Ordering Manager will provide tailored, intuitive screens and will be validated against both known data domains of the template and privileges of the product requestor. Feedback to users is provided on-line if those checks find discrepancies before a request is accepted. The same basic sequence is used to cancel or update orders. When a valid request has been submitted, the Common Ordering Manager assigns an internal order tracking number and prepares an electronic order request.

KMI-aware devices incorporate capabilities to automatically and directly interact with the Target KMI. In this regard, they can initiate KMI requests automatically, interacting with the PRSN Device Ordering Manager function in a fashion similar to the process used by authorized KMI users and managers. They will have to be registered as a valid end-entities and enrolled to authorize appropriate KMI privileges. Since there is no operator in the loop, they will not go through screens, but rather generate requests in an automated fashion. Orders from devices will be tracked in a standard manner so that device sponsor or component administrator can query status and intercede to update or cancel orders generated by devices under their purview.

## KMI Product Generation

In general, all KMI products will be generated within a PSN in response to order requests from a PRSN. These can result from product requests from KMI managers, directly from KMI-aware devices, or from event services. PSNs produce all physical KMI products. For electronic products, PSNs will only provide BLACK materials. The PSN will perform all cryptographic functions necessary to generate KMI products, and to protect them while being processed and stored within the PRSN, and for distribution directly to an end entity or through an intermediary (such as a Communications Security (COMSEC) Custodian).

## Delivery

PSNs arrange for delivery of all physical KMI products through proper physical distribution systems. However, the preferred distribution for KMI products is via Black, electronic transfers. The Target KMI is structured to enable delivery directly to end entities, including KMI-aware devices that can interact automatically with the KMI. This presumes that the KMI-

aware device include appropriate protocols to facilitate the transfers and internal cryptographic processing.

As discussed earlier, the PRSN Delivery Agent server can “push” products, automatically initiating an electronic transfer of Black KMI products over a secure link to a designated recipient. Authorized recipients can access the PRSN and log into the Delivery Agent server to “pull” KMI products. The Delivery Agent establishes a secure link with the intended recipient and electronically transfers the Black KMI products over that link. PRSNs include a capability for an electronic vault, providing a repository for previously generated and encrypted products, each with a unique identifier, split into a non-sensitive portion that is stored, and a sensitive portion that is encrypted. The PRSN is capable of querying to determine the status of materials that are stored, deleting stored materials, and retrieving them. If, to facilitate delivery, KMI products have to be decrypted (e.g., to make additional copies that can be prepared for delivery to multiple end entities), they are transferred back to a PSN for additional processing, and the requisite Black products are returned to the vault.

## Key Recovery

The KMI Key Recovery Agent capability will collect and archive all KMI information that may be needed to support key recovery operations. KRI will be encapsulated in a fashion to require multiple approved KRAs to collaborate to gain access the sensitive KRI. The encapsulation will enforce protection and access controls resultant KRI as dictated by appropriate national policies (e.g., two or more pre-selected individuals will have to be involved to gain access to the unprotected KRI materials.) When KRI is accessed, it will be protected to prevent inadvertent disclosure and transferred onto a KMI token for delivery.

## Revocation

Revocation of KMI privileges is accommodated by modification or deletion of roles and privileges as addressed under enrollment. The KMI will have the capability to revoke any KMI product. Each product will have a unique identifier (e.g., Certificate Number, Key Identifier). Authorized KMI managers can access the PRSN and log onto the Compromise Recovery Agent capability to process requests for revoking KMI products. When a validated request has been processed, the Compromise Recovery Agent will task an appropriate PSN to add the identified KMI materials to an appropriate mechanism to enforce the revocation.

The Target KMI will support two approaches for enforcing revocation. For certificate-based transactions, the KMI will integrate Online Certificate Status Protocol (OCSP) into their on-line validation servers. These servers provide worldwide distribution and access to information needed to ensure that only valid keys and certificates are being used. Protocols within KMI-enabled and KMI-aware applications and devices may include verification using these servers to show KMI materials at both ends of the transactions are valid. The Target KMI will also support the use of Compromise Recovery Lists including CRLs and CKLs as other mechanisms. These support other than certificate-based operations and are for use in situations where ready access to distributed, on-line servers is not operationally feasible (either based on mission constraints ala tactical environments) or at times when network access is limited or unavailable (e.g., network outages).

## System Management

Each KMI site has provisions for a site manager to perform a number of critical operations. A primary responsibility of these managers is to manage the day-to-day operations of the site. The site manager is responsible for monitoring the performance of the overall site, and when necessary off-loading operations to another site as a back-up capability. This includes a variety of tasks such as starting up, backing up, aborting, and restoring site operations. Other critical responsibilities are related to managing the security of the site, including the operation of intrusion detection systems (IDSs), local site responses to intrusions, managing of local security audits, sanitizing the site, and returning the site to a secure state. Site managers are responsible for coordinating the installation, testing, maintenance, configuration, and control of all components within the site.

The CSN is also responsible for the management of the overall KMI. In addition to its own site management, it provides long term archive capabilities, performs audits, provides help desk capabilities, and is responsible for enforcing and verifying compliance of operations with established security policies. The CSN is also responsible for management of all KMI IDS reporting, analysis of the aggregated information, and formulation and coordination of responses to suspected and actual cyber attacks.

Each PRSN site has several subsystems that provide data bases and data management services for the enclave. Each site will maintain, for example, the appropriate product catalog, registration data, and role and privilege data for clients that request products and services at that site. Each PRSN will also serve as a hot standby (back up) capability for other PRSNs and will have the capability for automated transfer of services to and from other PRSNs. Each PRSN also maintains a library of documents that can be downloaded by Client Node components, and software modules that may be run by clients that access the PRSN.

The PSNs also have system management responsibilities unique to their sites. As production nodes, they have to plan and schedule production activities (based on historical demand statistics and customer demand projections), monitor production flows, and allocate production resources to best satisfy production demands. To support tracing and status reporting of orders, the PSNs perform accounting and tracking of all orders from time of order receipt—through each production stage—until transfer of Black materials back to the PRSN or delivery of physical products directly to recipients. Since the PSNs process sensitive KMI product materials, they will have to maintain archive capabilities to augment those in the centralized long-term CSN archive, tools to facilitate appropriate audits, and facilities and procedures to comply with KMI security policies.

### 8.1.7.3.5 Transition

While the actual KMI structure will evolve over time, the KMI Program has established a fundamental philosophy for transition. Enhanced system capabilities will be introduced in parallel with existing operational capabilities. The strategy will be based on NO HARD CUTOVER whenever feasible. This will allow users to plan and implement effective transition of their operations to take advantage of new capabilities. Legacy capabilities will only be dismantled AFTER a complete operational transition has been accomplished.

## Impact of Transition on KMI Clients

Transition from the present systems to the Target KMI and the interim transitions from one KMI CI to another are planned and will be executed to minimize the impact on KMI managers and users. The Target KMI architecture itself has been designed consistent with this tenet. One example is the use of a “light client” concept to allow KMI manager workstations to remain stable, with enhancements being introduced in the servers typically provided in PRSNs. Another example is the use of validation servers to perform security-critical certificate path validation and enforcement of compromise recovery as a means for providing a more stable environment for client applications.

The PKI capabilities are intending to follow this to the fullest extent feasible. Adoption of commercial industry standards and trends are intended to maximize the use of commercially available applications. Reliance on commercial PKI tool kits for enabling of DoD custom applications is intended to ease PKI-enabling. However, commitment to commercial industry standards implies that custom DoD applications may have to be upgraded to follow the commercial sector’s evolution. If custom applications incorporate special features to support DoD-unique requirements, the diversity of COTS and GOTS systems can create significant issues.

Broader KMI capabilities will also continue to evolve. However, the KMI will maintain its full complement of products and services, and introduce new capabilities as additions, rather than replacements. As discussed above, KMI products and services will only be dismantled when the community no longer requires them. KMI tool kits will evolve to ensure backward compatibility as well as interoperability with the newest features of the KMI. KMI device owners, developers, and providers will have the opportunity to retain current operational configurations or take advantage of KMI advanced features, as they become available. The KMI’s longer range Capability Increment rollout planning enables device developers to plan their products’ evolution in an organized and efficient manner.

### 8.1.7.4 U.S. Federal Public Key Infrastructure

The Federal PKI is headed by the Federal PKI Steering Committee (SC) which is comprised of representatives from all Federal agencies either using or considering the use of interoperable public key technology in support of electronic transactions. The Federal PKI SC is chartered under the Enterprise Interoperability and Emerging Information Technology Committee of the US Federal Government Chief Information Officer (CIO) Council. It also has strong ties to the Security, Privacy and Critical Infrastructure Committee. It provides guidance to federal agencies and executive agents regarding the establishment of a Federal PKI and the associated services.

The Federal PKI SC also receives recommendations from the Federal PKI Technical Working Group (TWG), which responds to issues presented to it by the Federal PKI SC relating to the technical implications of developing the PKI.

The Federal PKI will support secure Federal Government use of information resources and the National Information Infrastructure (NII). The Federal PKI will establish the facilities, specifications, and policies needed by Federal departments and agencies to use public key



**UNCLASSIFIED**

Key Management Infrastructure/Public Key Infrastructure (KMI/PKI)  
IATF Release 3.0—September 2000

based certificates for information system security, electronic commerce, and secure communications.

The Federal PKI will support secure communications and commerce between Federal agencies, branches of the Federal Government, state, and local governments, business and the public. The Federal PKI will facilitate secure communications and information processing for Unclassified But Sensitive (UBS) applications.

The Federal PKI will be created largely from the bottom up. Federal efforts to use public key cryptography generally begin with individual applications within agencies that provide immediate support for vital agency programs. These implementations are paid for largely out of program funds, not funded as a centralized Government PKI.

The core Federal PKI consists of CAs, RAs, certificate status responders, and management authorities that manage public key certificates used by Federal departments and agencies for unclassified, sensitive applications.

PKI clients will use the public key certificates issued and managed by the PKI to provide security services to Federal users such as key pair generation, digital signature generation, digital signature verification, and confidentiality key management.

The Federal PKI is fielding a Bridge CA (BCA) that provides certification paths between CAs in agencies, and outside the Government. Federal CAs that meet the requirements of the Federal Bridge Certificate Policy will be eligible to cross-certify with the BCA, thereby gaining the certification paths needed for broad trust interoperability in the larger Federal and national PKI. Certificates issued to and from the Federal BCA will normally include certificate policy mapping extensions that allow relying parties to establish that remote certificate policies are equivalent to local ones. The Federal BCA operates under the control of the Federal PKI Steering Group, which is the Certificate Policy Authority for the Federal Government. Establishing policy mapping equivalencies is one of the Federal Policy Authority functions.

One of the drivers of the Federal Bridge CA design was the need to accommodate both the hierarchical and mesh PKI implementations that are already common within the Federal Government. Both hierarchical and mesh PKIs are operated by US Federal Government commercial and government partners. The BCA concept allows applications capable of processing mesh PKI certificates to interoperate with any mesh or hierarchical PKI cross-certified with the BCA.

Some commercial clients already include the certificate path development and validation capabilities needed to take advantage of the Bridge CA. Other vendors are now in the process of upgrading their PKI client applications with the features necessary to operate with the BCA. Figure 8.1-20 illustrates the planned architecture of the Federal PKI.<sup>5</sup>

The Bridge CA itself will actually consist of a variety of CA products that are mutually cross-certified. This design allows several vendors to operate within the Bridge CA “membrane,” thus allowing for continued BCA operation in the face of a dynamically changing PKI technology and vendor environment.

---

<sup>5</sup>Diagram courtesy of the Federal PKI web page at <http://csrc.nist.gov/pki/twg/welcome.html>.

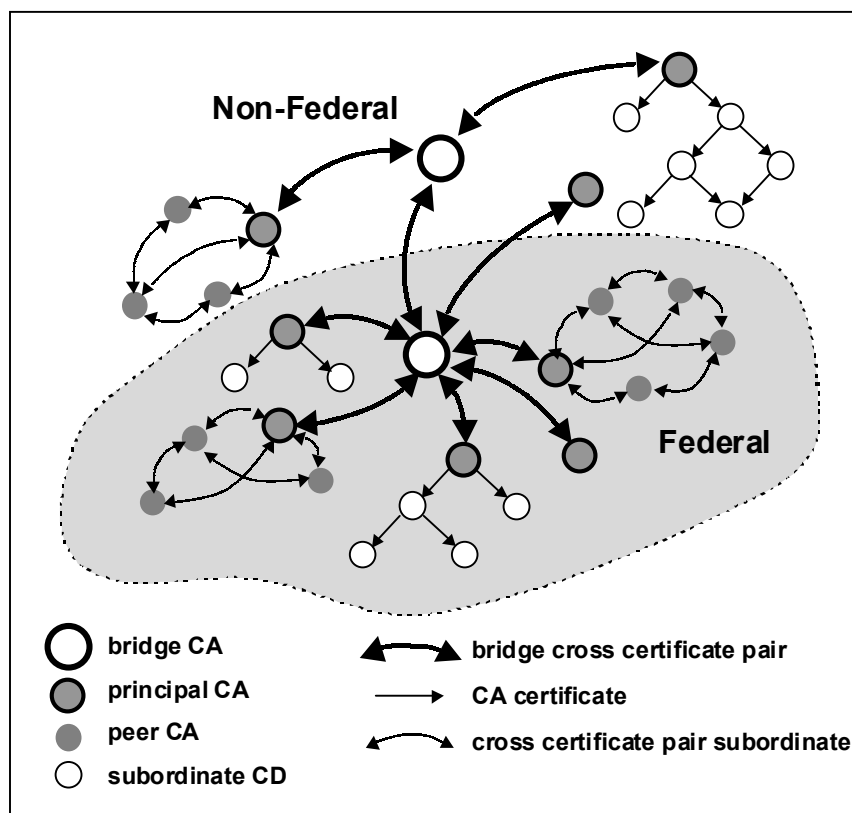


Figure 8.1-20 Federal PKI Architecture

## 8.1.7.5 Corporate PKI

### 8.1.7.5.1 Introduction

This section will focus on describing how the Microsoft Information Technology Group (ITG) built a public key infrastructure (PKI) by deploying a hierarchy of certification authorities (CAs) hosted on Microsoft Windows 2000 Servers. This name of this project was the Crypto Management Architecture PKI. For ease of discussion it will be shortened to CMA PKI.

### 8.1.7.5.2 Requirements

Microsoft is implementing and using many security technologies to protect and maintain the integrity of digital intellectual property. A large number of these security technologies depend on the use of valid X.509 certificates issued by trusted CAs.

The CMA PKI has to support the deployment of the technologies listed in the table below to satisfy the corresponding business requirements:

**Table 8.1-5 Business Requirement and Security Technology Comparison**

<b>Business Requirement</b>	<b>Security Technology</b>
Employees in all Microsoft business units need to exchange encrypted and/or digitally signed email with each other, external business partners, and customers over the Internet and other untrusted networks	Secure Multipurpose Internet Mail Extensions (S/MIME)
Secure networking with a common transport/tunnel technology supported by uniform authentication architecture	Internet Protocol Security (IPSec) and Layer 2 Tunneling Protocol (L2TP)
Users must be able to store encrypted data securely, while the corporation must be able to recover data should an employee leave or lose his/her encrypting certificate	Encrypting file system (EFS) and EFS recovery policies
Reduce the costs of purchasing certificates from outside sources by providing internally generated certificates for all intranet and most extranet SSL servers.	Secure Sockets Layer (SSL) or Transport Layer Security (TLS)
Strong authentication	Smart cards
Replace the practice of giving various external business partners shared corporate network accounts by trusting certificates from vendors/business partners	Certificates
Non-repudiation	Digital signatures

Additional requirements include:

- Active Directory integration - certificate revocation lists (CRLs), certificate enrollment, certificate templates, and CA certificates available via Active Directory
- Certificates mapped to users and computers in Active Directory
- Servers and client computers automatically enrolled for certificates (i.e. autoenrollment)
- Interoperability with Exchange Key Manager Server (KMS) and Outlook
- A healthy foundation for the expansion of Microsoft's corporate PKI to support forthcoming confidentiality, integrity, and authentication features in Microsoft products.

### 8.1.7.5.3 PKI Design

#### The Inherited PKI

At the beginning of the CMA PKI project, Microsoft already had a PKI managed by Legal and Corporate Affairs (LCA) and Product Release Services (PRS). This PKI was developed to support various product group and manufacturing efforts and was not utilized for general corporate functions.

Since Microsoft's root authority (MSROOT) in the inherited PKI is the top of the company's certification hierarchy for digitally signing all of its software products (among other things), a compromised MSROOT would have very negative national and global consequences. Therefore, the CAs that make up the inherited PKI are located in a secure vault on the Microsoft campus. The vault cannot be entered by a single individual, but rather must always be entered

## UNCLASSIFIED

Key Management Infrastructure/Public Key Infrastructure (KMI/PKI)  
IATF Release 3.0—September 2000

by two authorized individuals at the same time. The vault has also been designed to withstand attacks by cutting torches, explosives and other brute force tools of nefarious individuals.

## CMA PKI Topology

The CMA PKI has certification authorities in a three-level rooted hierarchy.

Level 1 - Microsoft Corporate Root Authority—The root certification authority at the top level of the hierarchy signs its own certificate. ITG makes it available to all entities that may want to establish trust in it.

Level 2 - Microsoft Intranet CA and Microsoft Extranet CA—The certification authorities below the root CA in a three-level hierarchy are referred to as a policy CAs or intermediate CAs. These CAs have certificates issued from the root CA and can be online or offline; ITG chose to keep the intermediate CAs offline for security reasons.

Level 3 - Microsoft Intranet CAs—The third level in a rooted hierarchy contains the issuing CAs. An issuing CA, as the name implies, issues certificates to end-entities. Issuing CAs are normally online CAs – in other words, they are always connected to the network.

## Certification Authority Servers

To establish the CMA PKI, eight certification authorities needed to be built. Three of the new CAs are off-line. These servers reside in the LCA vault.

The other five CAs will be on-line and service requests 24x7. These servers will reside in the ITG Vault.

## Microsoft Corporate Root Authority

The Microsoft Corporate Root Authority is a Windows 2000 CA. This represents the top of the Corporate PKI and is used only to sign/certify subordinate CAs. This server will be off-line except with generating revocation lists or signing CAs and will reside in the current LCA vault. This server should be built with the following parameters.

- Windows 2000 Certificate Server (Standalone Root CA).
- Self-signed CA certificate.
- Hardware-based Crypto Service Provider (CSP).
- 8 year CA lifetime.
- 2048 CA key length.
- 90 day CRL publishing interval.
- CRL Locations: Light-weight Directory Access Protocol (LDAP) to Active Directory; Hypertext Transfer Protocol (HTTP) to [crl.microsoft.com](http://crl.microsoft.com).

## Microsoft Intranet CA

The Microsoft Intranet Certification Authority will certify all other certification authorities used for internal purposes. This server will be off-line except with generating revocation lists or

signing certification authorities and will reside in the current LCA vault in. This server should be built with the following parameters.

- Windows 2000 CA (Standalone Subordinate CA).
- Install certificate from a PKCS#7 text file from the Microsoft Corporate Root Authority.
- Hardware-based CSP.
- 5 year CA lifetime.
- 2048 CA key length.
- 90 day CRL publishing interval.
- CRL Locations: LDAP to Active Directory; HTTP to [crl.microsoft.com](http://crl.microsoft.com).

### Microsoft Intranet Network CA

The Microsoft Intranet Network Certification Authority will issue end-entity certificates for services that relate to general server, user, or network administration such as Administrator certificates, EFS recovery certificates, router (IPSec/L2TP) certificates, and smart card enrollment agent certificates. The servers comprising this CA will be continuously on-line, require redundancy, and reside in the ITG vault. These servers should be built with the following parameters.

- Windows 2000 CA (Enterprise Subordinate CA).
- Install certificate from a PKCS#7 text file from the Microsoft Intranet CA.
- Hardware-based CSP.
- 2 year CA lifetime.
- 2048 CA key length.
- 24 hour CRL publishing interval.
- CRL Locations: LDAP to Active Directory; HTTP to [itgweb.corp.microsoft.com](http://itgweb.corp.microsoft.com).

### Microsoft Intranet FTE User CA

The Microsoft Intranet User Certification Authority will issue end-entity certificates to FTE users on the corporate network for general client authentication, EFS, and Smart Card Logon. The servers comprising this CA will be continuously on-line, require redundancy, and reside in the ITG vault. These servers should be built with the following parameters.

- Windows 2000 CA (Enterprise Subordinate CA).
- Install certificate from a PKCS#7 text file from the Microsoft Intranet CA.
- Hardware-based CSP.
- 2 year CA lifetime.
- 2048 CA key length.
- 24 hour CRL publishing interval.
- CRL Locations: LDAP to Active Directory; HTTP to [itgweb.corp.microsoft.com](http://itgweb.corp.microsoft.com).

## **Microsoft Intranet Non-FTE User CA**

The Microsoft Intranet User Certification Authority will issue end-entity certificates to non-FTE users on the corporate network for general client authentication, EFS, and Smart Card Logon. The servers comprising this CA will be continuously on-line, require redundancy, and reside in the ITG vault. These servers should be built with the following parameters.

- Windows 2000 CA (Enterprise Subordinate CA).
- Install certificate from a PKCS#7 text file from the Microsoft Intranet CA.
- Hardware-based CSP.
- 2 year CA lifetime.
- 2048 CA key length.
- 24 hour CRL publishing interval.
- CRL Locations: LDAP to Active Directory; HTTP to itgweb.corp.microsoft.com.

## **Microsoft Extranet CA**

The Microsoft Extranet Certification Authority will certify all other certification authorities used for external purposes. This server will be off-line except with generating certificate revocation lists or signing certification authorities and will reside in the current LCA vault. This server should be built with the following parameters.

- Windows 2000 CA (Standalone Subordinate CA).
- Install certificate from a PKCS#7 text file from the Microsoft Corporate Root Authority.
- Hardware-based CSP.
- 5 year CA lifetime.
- 2048 CA key length.
- 90 day CRL publishing interval.
- CRL Locations: LDAP to Active Directory; HTTP to [crl.microsoft.com](http://crl.microsoft.com).

## **Microsoft Personnel E-Mail CA**

The Microsoft Personnel E-Mail Certification Authority will issue end-entity certificates used for digitally signing and encrypting email (S/MIME). The server hosting this CA will be continuously on-line, require redundancy, and reside in the ITG vault. These servers should be built with the following parameters.

- Windows 2000 CA (Enterprise Subordinate CA).
- Install certificate from a PKCS#7 text file from the Microsoft Extranet CA.
- Hardware-based CSP.
- 2 year CA lifetime.
- 2048 CA key length.
- 24 hour CRL publishing interval.
- CRL Locations: LDAP to Active Directory; HTTP to [crl.microsoft.com](http://crl.microsoft.com).

## 8.1.7.6 Other Implementations

### The Kerberos Solution

Kerberos provides another approach for IA and network security. [7] Kerberos is a network authentication protocol. It is designed to provide strong authentication for client/server applications by using secret-key cryptography. A free implementation of this protocol is available from the Massachusetts Institute of Technology. Kerberos is available in many commercial products as well.

The Internet is an insecure place. Many of the protocols used in the Internet do not provide any security. Tools to “sniff” passwords off of the network are in common use by systems crackers. Thus, applications sending an unencrypted password over the network are extremely vulnerable. Worse yet, other client/server applications rely on the client program to be “honest” about the identity of the user who is using it. Other applications rely on the client to restrict its own activities with no additional enforcement by the server.

Some sites attempt to use firewalls to solve their network security problems. Unfortunately, firewalls assume that “the bad guys” are on the outside, which is often a very bad assumption. Insiders carry out most of the really damaging incidents of computer crime. Firewalls also have a significant disadvantage in that they restrict how your users can use the Internet. After all, firewalls are simply a less extreme example of the dictum that there is nothing more secure than a computer which is not connected to the network—and powered off! In many places, these restrictions are simply unrealistic and unacceptable.

Kerberos was created at Massachusetts Institute of Technology (MIT) as a solution to these network security problems. The Kerberos protocol uses strong cryptography so that a client can prove its identity to a server (and vice versa) across an insecure network connection. After a client and server have used Kerberos to prove their identity, they can also encrypt all of their communications to assure privacy and data integrity as they go about their business.

Kerberos is freely available from MIT, under a copyright permission notice very similar to the one used for the BSD operating and X11 Windowing system. MIT provides Kerberos in source form, so that anyone who wishes to use it may look over the code to assure themselves that the code is trustworthy. In addition, for those who prefer to rely on a professional supported product, Kerberos is available as a product from many different vendors.

In summary, Kerberos is another approach to your network security problems. It provides the tools of authentication and strong cryptography over the network to help secure your information systems across your entire enterprise.

### References about Kerberos

- More information about Kerberos can be found on the internet at <http://nii.isi.edu/info/kerberos>
- An excellent introductory article can be found at <http://www.isi.edu/gost/publications/kerberos-neuman-tso.html>

### 8.1.7.7 Additional References—Supporting Documentation on the Target KMI

As discussed in the roadmap, the Target KMI will be realized in an evolutionary fashion through a series of CIs. The definition of the Target KMI provides a perspective for each incremental capability to ensure that the goals established for it will be achieved. Specifically, the Target identifies the physical nodes, allocates the functionality within each node, and specifies the physical interface standards for KMI external and internal boundaries. The KMI Program has an ongoing system engineering activity to define and plan the target KMI definition. In January 2000, the KMI Program published a series of documents that describe the Target KMI definition that resulted from those activities. These documents include the following:

- a) KMI 2010 - Overview and Summary Information.
- b) KMI 2001 - Mission Needs Statement (MNS).
- c) KMI 2002 - KMI Operational Requirements Document (ORD).
- d) KMI 2000 - Functional Requirements Document.
- e) KMI 2022 - Standards and Technology Assessment.
- f) KMI 2020 - System Interface Description.
- g) KMI 2003 - KMI Security Policy and Requirements.
- h) KMI 2004 - KMI Threat Assessment Report.
- i) KMI 2005 - KMI System Security Architecture.
- j) KMI 2006 - KMI Security Risk Analysis/Assessment.
- k) KMI 2012 - Operational View (Concept of Operations).
- l) KMI 2011 - Program Glossary.
- m) KMI 2021 - Use Case Package (Five Volume Document).
- n) KMI 8000 - Target Architecture Validation Report.

### 8.1.8 Future Trends of Public Key Infrastructure

Public Key Infrastructure is one of the most promising technologies on the horizon today to provide strong authentication, data integrity, confidentiality, and non-repudiation services to a wide user base. The evolution of PKI has been dynamic and this trend will assuredly continue into the future. Although PKI products have been on the market for years, the technology still lacks maturity. Much work remains to be done by both product vendors and implementers. In addition, the public awareness of the benefits of PKI needs to be heightened before PKI will become the “silver bullet” it is intended to be.

One of the ongoing problems with PKI is incompatibility between vendor solutions. PKI standards need to continue to be developed and proven. While there are several major PKI vendors in the market place, many of the current products do not work with those from



another vendor. However, many vendors use the specifications provided by the RSA Security company, which have become in some cases, de facto standards.

There has been a growing trend toward standardization for certificates and cryptographic token storage formats. Through technical exchange meetings with vendors and standards groups such as the ETF and the ITU, it is likely that officially recognized standards will eventually be approved. These standards are vital for PKI to meet the demands for IA. PKI vendors have recognized this, and competing companies have shown increasing willingness to work together to produce common standards.

As better standards emerge, PKI products will improve. For example, the RSA PKCS #12 certificate container format allows private keys and certificates to be stored in a file on a disk. Access to this information can be protected by a password. Since the user chooses the password, a bad password choice can impair the security of the stored certificate. Despite its disadvantages, the PKCS #12 format is the most widely used format and at the present time, there is no widely available alternative that is suitable for replacing PKCS #12. An improved method is needed, but a new method needs to be accepted by the entire industry to become successful.

Vendors that produce interoperable products allow enterprises to purchase PKI equipment with less fear that the purchased product will become obsolete or will no longer be supported. Instead, it is known that the product will operate with others, even if they are produced by a competitor. When the time comes to upgrade, upgrades are less painful if mature standards are in place. The upgrade can be phased in over time, and there should be a richer set of upgrade features from which to choose. There is a wide variety of unrelated applications that could benefit from a common security solution. A common system reduces the long-term costs associated with maintaining a separate security infrastructure for each application. PKI could provide this solution, and interoperability among a common PKI is important.

Over time, the underlying cryptography of PKI will need to change continuously. It is obvious that new computers are constantly becoming faster. Faster computers will benefit the “brute force” method of cracking encrypted information. As such, the encryption technology must improve to stay ahead. As consumer computers are able to process data to crack the current encryption scheme in a reasonable period of time, data protected by cryptographic techniques becomes more endangered. Even without advances in computer speed, advances in other areas, such as distributed computing, will make encryption upgrades a requirement. A PKI integrator should not assume that a major investment in PKI would be a one-time expense.

### 8.1.8.1 Smart Cards

One of the promising new PKI implementations will be smart cards. Smart cards will provide vast new advantages for PKI. Private keys will be stored in a microchip on the card, rather than on a computer disk. The smart card contains not only data, but also a microprocessor to manipulate and protect the stored data. The smart card can control access to private key on the card, and prevent unauthorized manipulation of the data.

Once the private key has been generated by a smart card, the on-board crypto-processor contains the private key. This processor prevents outside access to the private key. Smart cards

also offer an alternative to the limitations of the RSA PKCS #12 certificate container, by providing additional security to the private key.

Smart cards will provide mobility to PKI users. A single card could be used for physical access to a building, to login to a computer, and to securely transmit information.

There are some disadvantages to smart cards. An obvious disadvantage is that they might easily become lost or stolen. While a stolen smart card should not reveal any information to its finder, its legitimate owner may not have a means to access his computer system, or gain access to a building. Another disadvantage of smart cards is that it may be desirable to operate several computer systems each of which employ a smart card. If a user has only one smart card, or does not have enough to use simultaneously, the smart cards will not be useful.

## 8.1.8.2 Biometrics

Biometric devices represent another emerging technology. These devices use physical features or behavior characteristics of human beings to identify a person. Biometric devices will measure unique qualities, such as of a person's retina or fingerprint. Upon login, the devices measure the appropriate qualities of the user. The devices then compare those qualities with known qualities, which are stored digitally.

The technology is advancing rapidly. When combined with PKI and smart cards, biometrics offer additional advantages. PKI alone cannot guarantee the identity of a person. The person using the PKI usually enters a password or personal identification number to access her private key, and identify herself to the PKI. If this password is compromised, the PKI is compromised. Instead of using a password, a user could use a biometric device to authenticate herself to a PKI system via a biometric device. The biometric device provides additional assurance that the person is actually who she claims to be. The addition of biometrics is a solution when assurance of authentication to the PKI is essential.

At the present time, well-chosen and protected passwords can provide a higher level of assurance than biometric devices, due to their lower probability of being guessed versus the higher probability of a biometric device mistakenly identifying a person. As biometric devices improve, their accuracy is likely to improve significantly. Biometric devices offer increased value by taking some risks out of user passwords. Examples of password risk include users choosing simple, easily guessed passwords, or users writing passwords on a piece of paper that is not properly secured.

Biometrics are expected to grow significantly in the security field within the next ten years. Although prices are still relatively high, they will come down in time. Several companies are already marketing biometric devices to the public. The combination of biometrics with PKI provides synergy between these two technologies. Biometrics provide a more secure login than a simple password access to one's private key, and PKI allows biometric devices to be used across a wide system infrastructure. Disadvantages to biometrics include the user's resistance to biometrics requirement that their personal qualities (e.g., retina image) be examined or stored, and the relatively high cost of the biometric devices.

### 8.1.8.3 Certificate Revocation

A certificate revocation scheme needs to be in place to prevent a user's certificates from being valid when a PKI user has their access to the PKI removed. For example, an employee who leaves a position or is transferred to another position will likely need to have access removed. Since this user's public key may still exist in the local directories of other user's computers, there needs to be a method in place to prevent the certificate from being used. There are two leading methods being investigated to accomplish this: CRLs and the OCSP.

CRLs are a comprehensive record of all certificates that have been previously issued but are no longer valid. The CA publishes, and is responsible for, the CRL. The CRL includes the serial numbers of all certificates that have been revoked. This scheme requires that a client wishing to check a certificate against a CRL to download the entire CRL. The CRL would then be searched to discover if any listed certificates match the certificate that the client is checking. An expiration date is included in the CRL, at which time the CRL must no longer be relied upon for validation.

The OCSP is another method to ensure the currency of a certificate. A work in progress by the IETF, it employs a client/server approach. A client wishing to validate a certificate sends a request to a server. The request includes a list of certificates or serial numbers that the client wishes to check. The server sends back a reply, which is signed by a CA to ensure the validity of thereply. The reply has several possible responses: Not Revoked, Revoked, On Hold, or Expired.

At this time, there is no clear consensus on which method will prevail. Certificate revocation schemes will be a major task of future PKI development.

### 8.1.8.4 Certificate Recovery

A key recovery system might be employed on some PKIs. The recovery system allows access to the private key through an alternate means. For example, this is useful if the user forgets a password, or management must know the contents of a user's encrypted message. Key recovery systems may be appropriate for encryption keys, but are not recommended for identity keys.

Identity keys are used for identity purposes only. For example, when a user wishes to add non-repudiation benefits to an email message, he can sign the email with his private identity key. An encryption key is used to provide confidentiality services. If the user wishes to send a confidential email message, the public encryption key of the addressee would be used. The private key of the addressee is required to view the contents of the message.

A key recovery system will allow the encrypted data to be made available to the trustees of the key recovery system. Since an identity key is only used to provide identity services, there is no legitimate reason to recover the key. If the password to the identity key is lost, the key can be revoked and a new key issued. A PKI policy can help prevent misuse of identity keys to falsely impersonate a user by not permitting identity keys to be escrowed in a key recovery system.

Key recovery systems have serious security ramifications. Introducing a key recovery system into a PKI introduces a weak link in the security chain. While key recovery systems can be a method to help guard against dishonest users, there is no guarantee that a person entrusted

## UNCLASSIFIED

Key Management Infrastructure/Public Key Infrastructure (KMI/PKI)  
IATF Release 3.0—September 2000

with the key recovery system will not be dishonest as well. A breach of security in this system could remove virtually all of the security advantages of PKI. In the future, biometric devices might help prevent the lost password problem. If key recovery systems are still desired for other reasons, they should be employed with great care.

### 8.1.8.5 KMI

The KMI is a common structure to administer keying material within DoD. The KMI will eventually administer all keying material throughout DoD. This material includes legacy symmetric key products, and public (asymmetrical) key products. As KMI becomes a common administration tool for all DoD keys, it will be used for key registration, key generation, secure key archiving, and key distribution. Additional systems are being examined to study the feasibility of integrated into the KMI.

The KMI architecture will likely consist of several nodes. A CSN will provide data storage, a root certificate authority, archive audit records, and intrusion detection systems. A PSN will provide key generation services and certificate generation at the certificate authority level. A PRSN will provide key registration, tracking, directory services, key recovery services, and privilege assignment. The Clients Node will distribute keys, and provide an interface for customer services.

The CSN is envisioned to have KMI databases and library services. It would provide support to supervise the KMI system, and could operate at the SECRET level. The PSN will likely be designed with a modular construction. The key generation and management functions can be added or deleted as they are needed. The PSN would support new services, as they become available. The PRSN would be deployed on the DoD networks (e.g. NIPRNET, SIPRNET) and would be intended for deployment on a regional basis. The PRSN would operate at the network's classification level. It would provide support for key recovery services within the KMI.

The KMI will likely need to be accredited to operate at system high. Various nodes will operate at TOP SECRET-high, SECRET-high, etc. as needed. Provisions will be need to in place to isolate nodes with dissimilar classifications, and to prevent data cascading to a lower classification. In the future, it is possible that the DoD KMI will interface with other KMIs within the U.S. and its Allies. Policies will need to be changed to allow crypto data transmission over protected Local Area Networks (LANs) such as SIPRNET.

A KMI and a PKI are closely related technologies, and are designed to work together. The KMI will provide support for the keys that the PKI must use. The PKI program benefits by making use of an existing key infrastructure, while providing new capabilities. According to the NSA KMI Standards and Technology Survey, key management will be accomplished in a similar method to that developed for multicast groups. Policies are constructed for a number of groups. Group keys are created by a group controller, which then distributes them. The Group Secure Association Key Management Protocol (GSAKMP) is then used to distribute the groups' policies and provide for future re-keying of each group when needed.

The KMI is currently a work in progress. The plans for the system will likely change as it is designed and built. It is uncertain at this time how the KMI will be modified or what additional users it will eventually serve.

### 8.1.8.6 Risks Associated with this Analysis

This analysis of what PKI will be like in the future consists of predictions based upon current trends today. The PKI momentum has been building for several years and this is likely to continue. However, PKI has shown fairly slow growth so far. The growth is not widespread at this time outside of a few select industries. As standards and new technologies mature, PKI will likely become much more important.

There are several risks in predicting the future trends of PKI. Usability will be an extremely important factor in that maturation of PKI. While important advances in this area have been made, more will need to take place in the future. It is also possible that another technology will emerge that can provide similar benefits and will be more efficient to deploy. At this time, the future of asymmetric keys to provide strong authentication, data integrity, confidentiality, and non-repudiation services appears to be solid. PKI is the technology most likely to benefit from the advantages of asymmetric keys to provide these services.

### 8.1.8.7 Conclusions

PKI can be expected to grow vigorously in the next five to ten years. As standards are developed and more applications are supplied with PKI built in, the PKI will grow more quickly. It is possible that one or more competing technologies will arise on the security scene as well, but such a technology will likely provide similar capabilities that PKI promises. The advantages of PKI will be the flexibility to adapt to new applications, and to provide a common security architecture that can be deployed for many applications, involving both computers and other devices.

The future of PKI will largely depend on its usability. Even the best security resources cannot provide security if they are not accepted by end users. PKI offers a number of benefits and is intended to be used for more than one application. For example, an email system may use PKI for confidentiality and non-repudiation across an enterprise and to operate with external enterprises. A database system might use the same PKI to provide confidentiality and non-repudiation plus authentication to the database as well. As more applications use a common PKI, additional economies of scale can be realized. Existing applications will need to be replaced with newer software that is PKI compliant, or PKI enabled. Application integration will likely be one of the most difficult and most expensive phases of adopting a common PKI system.

## References

---

1. International Telecommunication Union (ITU), 1997, Information Technology - Open Systems Interconnection - The Directory: Authentication Framework, ITU-T Recommendation X.509.
2. RSA Laboratories, November 1, 1993, PKCS#10: Certification Request Syntax Standard, Version 1.0.
3. RSA Laboratories, November 1, 1993, PKCS#7: Cryptographic Message Syntax Standard, Version 1.5.
4. Bruce Schneier. Applied Cryptography pp139-152
5. PKIX -4. Public-Key Infrastructure (X.509) (pkix) 7 August 2000  
<<http://www.ietf.org/html.charters/pkix-charter.html>>.
6. PKI Profile. NIST PKI Program 23 February 2000 <<http://csrc.nist.gov/pki>>.
7. Massachusetts Institute of Technology's Kerberos: The Network Authentication Protocol Web Site, June 24, 2000, <<http://web.mit.edu/kerberos/www>>.

## Additional References

---

- a. Furlong, Judith, Public Key Infrastructure (PKI) Scenarios Overview, 20 November 1997.
- b. University of Southern California The Kerberos Network Authentication Service  
<<http://nii.isi.edu/info/kerberos>>.
- c. B. Clifford Neuman and Theodore Ts'o. Kerberos: An Authentication Service for Computer Networks USC/ISI Technical Report number ISI/RS-94-399. September 1994  
<<http://nii.isi.edu/publications/kerberos-neuman-tso.html>>.
- d. The Moron's Guide to Kerberos, Version 1.2.2,  
<<http://www.isi.edu/gost/brian/security/kerberos.html>>.

**UNCLASSIFIED**

Key Management Infrastructure/Public Key Infrastructure (KMI/PKI)  
IATF Release 3.0—September 2000

**This page intentionally left blank**

**UNCLASSIFIED**

8.1-103

## 8.2 Detect and Respond as a Supporting Element

A fundamental tenet of the defense in depth strategy embraced by this Framework is to prevent cyber attacks from penetrating networks, and to detect and respond effectively to mitigate the effects of attacks that do. An integral aspect of this strategy is a secure infrastructure to support the detection of and reaction to cyber incidents and attacks.

### 8.2.1 What This Focus Area Addresses

Detect and Respond capabilities are complex structures that run the gamut of intrusion and attack detection, characterization, and response. The progression of Detect and Respond technologies is building from audit logs and virus scanners to a more robust capability. While technology advances continue to evolve, this overall area remains heavily dependent on highly skilled operators and analysts.

#### 8.2.1.1 Scope of This Focus Area

The local environments (within an enclave) are the logical location for network-based and host-based sensors. Sections 6.4 (Network Monitoring within Enclave Boundaries and External Connections), 6.5 (Network Scanners within Enclave Boundaries), and 7.2 (Host-Based Detect and Respond Capabilities within Computing Environment) address specific Framework guidance for these sensors. This section addresses the processes and technologies that are typically required beyond the sensors. This includes discussions of architectural considerations for improving the Detect and Respond posture of an enterprise, evolving paradigms for a Detect and Respond infrastructure, the various processes and functions that are performed within the secure infrastructure, and the technologies that are available to realize these processes and functions. The section concludes with sources for additional information and a list of references used in developing this guidance.

#### 8.2.1.2 Terminology

To set the stage for the discussions in this section of the Framework, there are a number of terms that should first be defined. We recognize that these terms, which are fundamental to the discussions in this section, are also germane to many sections of the Framework. We also appreciate that these terms have varying interpretations within the community, so we include the following definitions to eliminate possible confusion or ambiguity within this section of the Framework.

The first set of terms deals with threats and vulnerabilities. A *threat* exists when an intruder (also referred to as an *adversary* or a *threat agent*) has the means, motivation, and opportunity to exploit an information system and/or its associated networks. A *vulnerability* is a weakness or



hole that can be exploited by an intruder. An *attack* is a sequence of events an intruder uses to exploit a vulnerability.

An intrusion can be thought of as an attempt or actual break-in into an information system. The intruder's intent may to misuse the system or data contained within the system, render a system unreliable or unusable, gain access to the data contained on the system, and/or manipulate the data. Once an intrusion has occurred on an information system, the damage can be extensive—sensitive information may be compromised and network systems or network services can be rendered inoperable. These events can result in the loss of a corporation's competitive edge, lost productivity when network services are unavailable, and costly man-hours and dollars to assess the impact of an intrusion and recover any lost data.

Beyond this, there are various levels of an “attack” that are also worth identifying. We look at attacks from a bottom's up perspective, since they are detected based on a logical progression from the point of view of sensors (e.g., intrusion detection system or IDS).

- *Alarms* are the typical output provided by a sensor as an indication that it believes it detected some evidence of the presence of an intruder.
- *Events* are actual occurrences of some irregularity that caused an alarm. We distinguish alarms from events in that there are often a number of valid network and host operations that may cause an alarm (thus giving rise to false positive indications).
- *Interesting Events* are based on the recognition that local environments may experience hundreds of thousands of events daily, and there are typically only a small number that have the potential for any real damage. This category represents those that have the potential for serious impact such as may be characterized in a security policy.
- *Incidents* are interesting events that actually have serious impact on the information systems and networks of a local environment.
- *Attacks* are concentrated efforts by an adversary or intruder to have serious impacts to an overall enterprise, usually implemented by a series of incidents targeted at multiple local environments.

While all incidents and attacks are important, the Framework guidance focuses on attacks where the attacker(s) have the will, resources, and persistence to cause grave harm to an enterprise.

## 8.2.2 Enterprise Architecture Considerations

While planning for a Detect and Respond infrastructure, it is important to recognize that the enterprise networks and systems that it will support must also be structured to provide information to, and take advantage of the services and information such a secure infrastructure provides. The remainder of this section provides guidance on configuring an enterprise to improve its Detect and Respond posture.

### Incident Reporting

As highlighted in Sections 6.4 (Network Monitoring within Enclave Boundaries and External Connections), 6.5 (Network Scanners within Enclave Boundaries), and 7.2 (Host-Based Detect

and Respond Capabilities within Computing Environment) of the Framework, the local environments have the option of deploying sensors, and possibly analysts to interpret the results of, and, when appropriate, react to the implications of these outputs. Beyond the local environment, each organization, or perhaps community has to determine what information should be reported, in what format, under what situations, and to whom. The Department of Defense (DoD) has issued implementation guidance and a joint policy for incident and vulnerability reporting. Other system infrastructures simply allow reporting, and leave it to the local environment to work directly with the next tier to decide when, what, and how to report.

## Network Partitioning and Redundancy, Back-up

Networks are typically configured to provide the most cost-effective service to its users. Whenever feasible, networks should be partitioned into logical segments, with boundary protection devices between segments. This limits traffic flow and thus potential exposure within segments, provides a degree of isolation if one segment or another is subverted, and facilitates the shutting down or limiting of services within effected segments as a possible response. Offering redundant capabilities within a network creates the potential for response options where authorized traffic can be diverted around a segment that has been exploited.

## Deploy Technical Safeguards and Countermeasures as Response Options

A fundamental aspect of an effective React capability is to deploy safeguards and countermeasures that can be activated to implement responses. Whether they are making changes to firewall policies, filtering router configurations, deception servers, or others, there are a number of such countermeasures available, as discussed in Section 8.2.5.4 (Response Tools).

## Plan for Contingency Operations

There is an entire discipline associated with disaster planning (sometimes referred to as planning for contingency operations) that includes the development of anticipatory processes and procedures that can facilitate an effective response. These include creating back-ups of mission-critical and establishing pre-planned courses of action (COAs). Recommendations regarding the preparation of COAs include the following:

- Plan to deal with high probability threats and at least acknowledge the less likely possibilities.
- Allocate resources to complete and coordinate the planning; create plans in advance rather than waiting for an event to occur.
- Coordinate and obtain approval/acceptance of plans by upper management, business unit managers, and other decision makers.
- Take advantage of planning that other, similar organizations may have already prepared.
- After the plans are formulated, exercise the procedures to validate the approach, refine the tactics, and train the participants.

When the program is in place, frequently review, update and enhance it to keep it current.

## Coordinating Responses

Fundamentally, response itself is an issue for the local environments. However, there are a number of factors with implications beyond the perspective of local sites that need to be considered when formulating and evaluating response options, as well as when actually responding to an intrusion or attack. A basic decision is whether to shut down an intruder's access (or an entire site) or to allow an intrusion to continue while evidence that will be needed for subsequent prosecution is collected.

## Considerations for Operations

As with the architectural features identified above, there are also complementary operational practices<sup>1</sup> that are important to the overall defense of an enterprise, and again, are directly relevant to considerations for a Detect and Respond infrastructure:

- Be prepared for severe denial-of-service attacks (e.g., institute and practice contingency plans for alternate services).
- Inspect for physical penetrations.
- Educate users and staff.
- Institute well-known procedures for problem reporting and handling.
- Institute procedures for reporting suspicious behavior.
- Institute and monitor critical access controls (e.g., restrict changeable passwords, require dial-back modems).
- Minimize use of the Internet for mission or time critical connectivity.
- Require security-critical transactions (e.g., establishing identity when registering) to be conducted in-person.
- Institute and monitor a strict computer emergency response team alert and bulletin awareness and patch program.
- Establish procedures for recovery from attack.

## 8.2.3 General Considerations for a Detect and Respond Solution

It appears that there are no generally accepted architectural constructs for a Detect and Respond infrastructure across various communities. However, there are several fundamental considerations for a Detect and Respond infrastructure that appear to be consistent across communities. These are highlighted below.

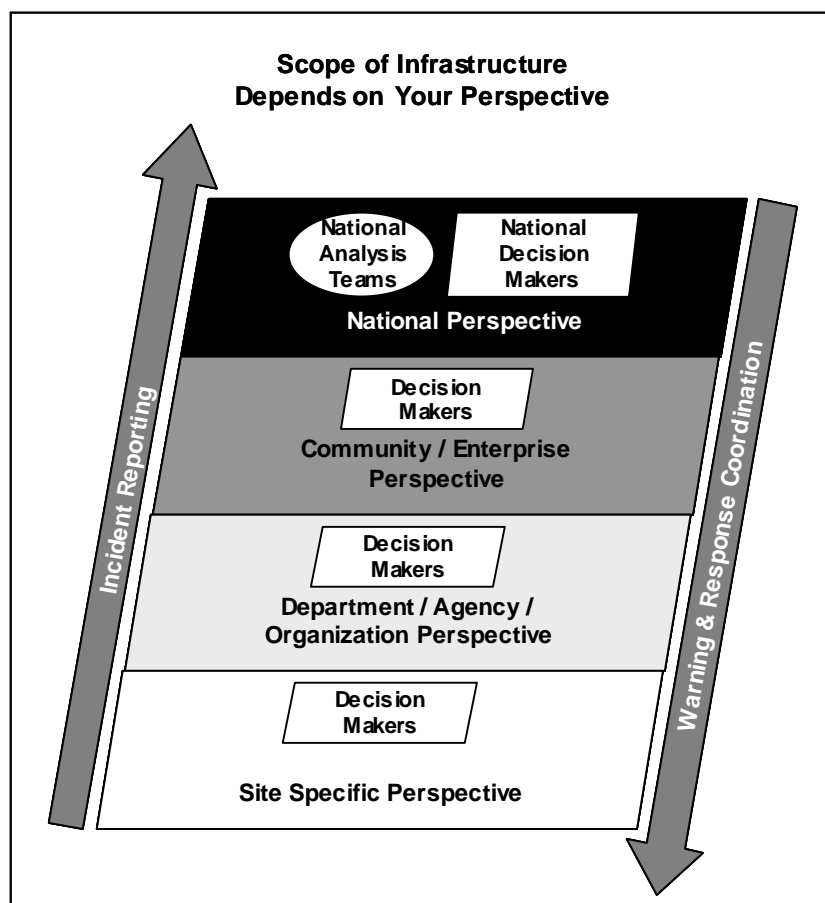
---

<sup>1</sup>Note that it is imperative to perform quality network management and system security administration in order to maximize the security of the network configuration and mechanisms and to increase the likelihood of detecting and successfully reacting to attacks.

### 8.2.3.1 General Constructs for a Detect and Respond Infrastructure

In general, many network infrastructures are inherently hierarchical by their nature, and this one is no exception. When considering a general construct for a Detect and Respond infrastructure, a primary consideration is the perspective that the system infrastructure layer will maintain for its support. Figure 8.2-1 identifies typical layers in this hierarchy, and the perspectives that each layer could offer. Each layer usually retains responsibility for its own operation, and thus must be capable of making decisions about courses of action for its own operation. However, it is seldom the case that any site can function in a completely autonomous fashion without some oversight, coordination, and direction, so there is a natural hierarchy for the decision making as well.

What tends to be the case in general is that information about incidents, which is usually sensed at the lowest layer in the hierarchy, is promulgated up to higher layers with some form of reporting. Warning and response coordination that are more typically derived from higher layers are disseminated from these higher layers down. Again, these are general statements, and any specific situation has to be tailored to the unique needs of the constituent segments.



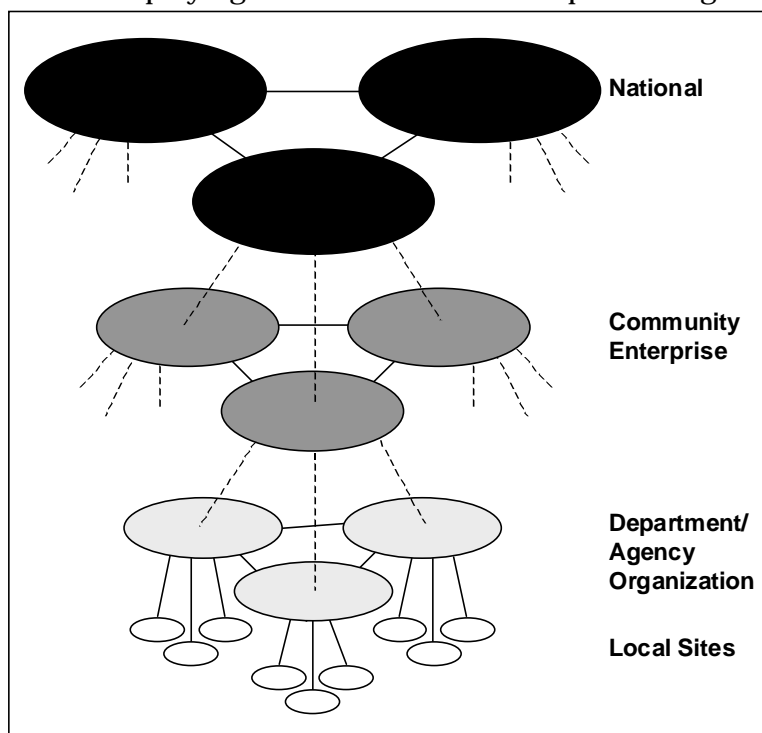
**Figure 8.2-1 Perspectives of Layers in a Detect and Respond Infrastructure Hierarchy**

### 8.2.3.2 Examples of Existing Detect and Respond Infrastructures

A Detect and Respond infrastructure of this nature will likely be structured in the manner depicted in Figure 8.2-2. This is reminiscent of, and consistent with various actual hierarchy

structures used today in various communities and enterprises. The specific relationships and responsibilities across the layers differ in actual practice.

For the DoD, local sites are responsible for deploying network monitors and performing site assessments. Typically each Military Department (MILDEP) has its own Navy Computer Emergency Response Team (NAVCERT) capability or Air Force Information Warfare Center (AFIWC) that is responsible for attack detection and characterization for that MILDEP. At the Enterprise level, DoD has established a Joint Task Force for Computer Network Defense (JTF-CND), with a technical analysis capability within the Global Network Operations Security Center (GNOSC) to monitor critical defense networks and coordinate actions across the DoD to restore functionality after an intrusion or attack. The DoD model differs from the others in that reporting and response coordination procedures are mandated.<sup>2</sup>



**Figure 8.2-2 Basic Hierarchy for Detect and Respond Infrastructure**

The Civil Government agencies have adopted a less formal structure. There is a Federal Computer Emergency Response Team (FEDCERT) that is responsible for coordinating Detect and Respond activities across the Federal Government, but its use appears to be at the discretion of individual agencies. Selected agencies maintain their own Computer Emergency Response Team (CERT) capabilities (e.g., Department of Energy (DOE) Computer Incident Advisory Capability, or CIAC that is operated at Lawrence Livermore National Laboratories as a central clearing house for reporting incidents.) This community also takes some advantage of CERT capabilities from academia (e.g., CERT associated with Carnegie Mellon University actually funded by DoD). The Federal Intrusion Detection Network (FIDNet), a General Services Administration (GSA) initiative to centralize a Federal Government-wide capability to analyze local sensor outputs is consistent with this general hierarchy, but may be implemented as a managed commercial security service offering available to those agencies that decide to subscribe.

In the private sector there are CERTs available to support those specific organizations that choose to use them, again with reporting and coordination at the discretion of the organization.

<sup>2</sup>The Department of Defense has issued CJCSI 6510.01B, a JCS publication providing implementation guidance and a joint policy for Defensive Information Operations. Within that document, Enclosure D, Appendix G defines incident and vulnerability reporting procedures, methods, and reporting formats.

The Information Sharing and Analysis Center (ISAC), a construct resulting from efforts to implement Presidential Decision Directive PDD-63 was conceived as a mechanism to structure “sector” (e.g., banking and finance, telecommunications) coordinators. The intent was to provide a mechanism for enabling appropriate, anonymous, and confidential sharing of information associated with incidents, threats, vulnerabilities, and solutions associated with each sector’s critical system infrastructures and technologies. One ISAC is in place for the banking and finance community. While others have not been put into operation, it is again representative of the use of a hierarchical structure for a Detect and Respond infrastructure.

At the national level, the National Information Protection Center (NIPC), established at the FBI again in response to PDD-63, is intended to serve as the U.S. Government focal point for threat assessment, warning, investigation, and response to threats or attacks against our nation’s critical infrastructures. They are supported by the National Security Incident Response Center (NSIRC) at National Security Agency (NSA) to bring perspectives from the Intelligence Community to perform in-depth analysis (including post-attack investigation) to support activities at the NIPC (and JTF-CND). While these national layers of the infrastructure are called upon at the discretion of other organizations, they maintain a national level perspective. The NIPC also leads or coordinates activities associated with national security or criminal investigations of cyber crimes.

Although not depicted in the figure, there is some evidence of global infrastructures being established at the international level. One such example is the Forum of Incident Response and Security Teams (FIRST), whose membership includes DoD Service CERTs, academia, and major private corporations from across the globe. Their goals are to foster cooperation among constituents for the protection, detection, and response from computer intrusions. They provide a means for sharing alert and advisory information, and facilitate collaborative planning and sharing of information, tools, and techniques.

## 8.2.4 Detect and Respond Functions

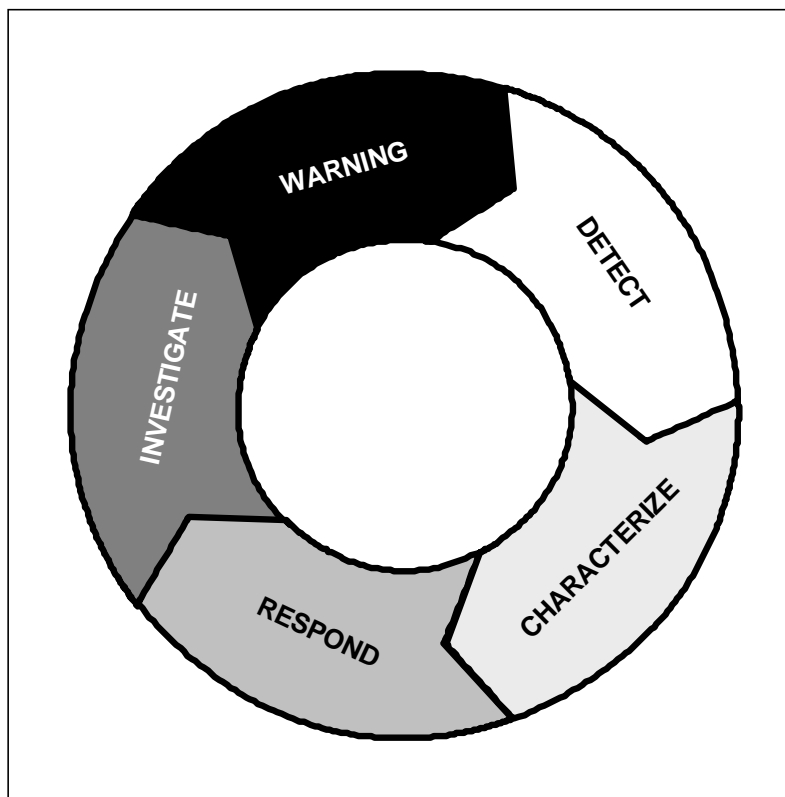
Looking across the existing Detect and Respond infrastructures, there is a general recognition that a wide range of functions is needed to support their operations. In many cases, technology solutions are not available to perform these functions automatically. Analysts, network operators, and system administrators perform many of the functions by applying basic support technologies to ease their tasks. This section provides an overview of the functions that these analysts (with their tools) are attempting to perform. This section begins with an overview of the various phases of operation associated with Detect and Respond and then highlights specific functions that are representative of each phase. The section that follows provides a discussion of the underlying technologies that are available to support Detect and Respond capabilities.

### 8.2.4.1 Phases of Operation

There are five basic phases of Detect and Respond that should be considered:

- Warning – Providing advanced notice of a possible impending attack, including a perspective on the attack strategy, scenarios, likely target sites, and timing.

- Detect – Determining that an attack is, or has occurred. This includes the sensing functions discussed in Sections 6.4 (Network Monitoring within Enclave Boundaries and External Connections), 6.5 (Network Scanners within Enclave Boundaries), and 7.2 (Host-Based Detect and Respond Capabilities within Computing Environment) of the Framework, along with broader activities to discern an attack (versus an incident) is underway



**Figure 8.2-3 Basic View of Detect and Respond Phases**

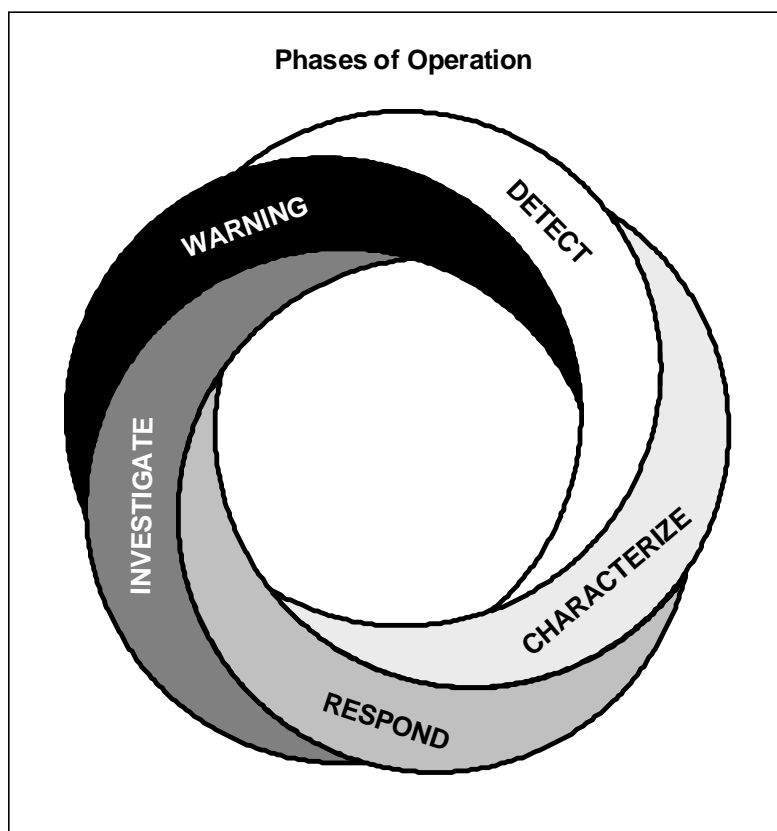
- Characterize – Analyzing the attack in term of its intent, approach, projections of how it will proceed, likely impacts, and possible identification of the attack source
- Respond – Reacting to mitigate the effects of the attack and restoration to return the systems and network to normal operation
- Investigate – Analyzing how an attack was accomplished in order to provide feedback to improve existing Protect, Detect, and React capabilities to ensure that similar exploitations cannot occur, and when appropriate to provide evidence when prosecution of attackers is pursued.

From a process standpoint, it is possible to consider Detect and Respond operations as a series of phases, or stages that in essence form a life cycle for a particular incident or attack. Figure 8.2-3 provides a high level view of these phases. In this view, it is easy to consider the cycle of phases to begin anew with the occurrence of another attack.

While this perspective is straightforward, it is not really reflective of real-life situations. While there is sense of “hand-off” from one phase to another, each of the phases is really an ongoing set of processes. For example, warning does not typically stop after an alert is issued.

It continues to search for new indications while detection capabilities focus on those being anticipated. This is typically the same for each phase, as represented in Figure 8.2-4. This sort of twisting view of Detect and Respond phases may seem whimsical, but is really more indicative of practical operations.

There are a number of approaches for realizing these phases within the context of a Detect and Respond hierarchy. Figure 8.2-5 provides a perspective that can be used when considering allocation of Detect and Respond functions. While each local site, organization, or enterprise (community) has the option for allocating Detect and Respond functions within their hierarchy, it is often the case that Warning and Attack Investigation are provided as Detect and Respond infrastructure services since they require highly skilled analysts and access to broad and diverse sources of information. The other functions tend to follow the perspective on the hierarchy level. Thus, the functions on the left side of the figure that focus on incidents are typical of those of a local environment, or possibly an organizational level. Those on the right side of the diagram that focus on attacks (versus incidents) are more indicative of those of a higher level of the system infrastructure (based on the view that attacks are really comprised of coordinated incidents across multiple sites.)



**Figure 8.2-4 Realistic View of Detect and Respond Phases**

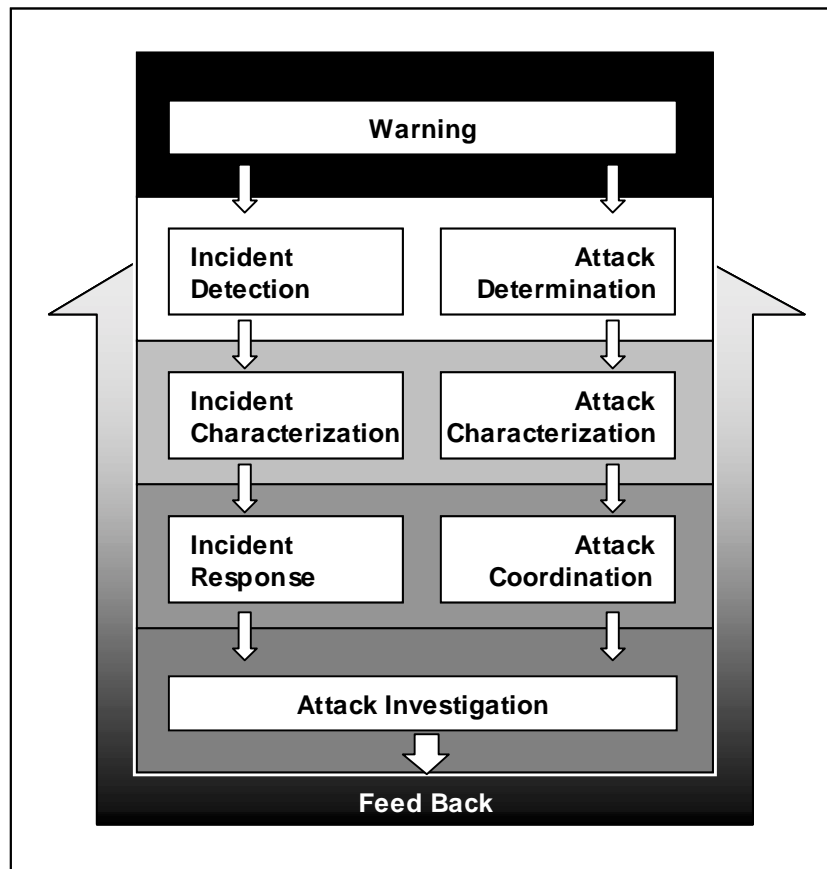
Another important aspect of these functions is that they are highly dependent on one another. They each rely on, and provide information to others, working toward a common goal of successful detection and response to incidents and attacks. The following section highlights representative processes for each of the eight functions identified in the figure. Again, these are offered not as direction of what functions have to be performed, but to offer a perspective on what detection and response must achieve using the available technologies discussed in subsequent sections.



## 8.2.4.2 Functions to Support Warning

Warning is a proactive capability intended to provide advanced notice (or warning) of possible impending cyber attacks. Figure 8.2-6 offers a perspective on the types of functions that could be implemented to support warning.

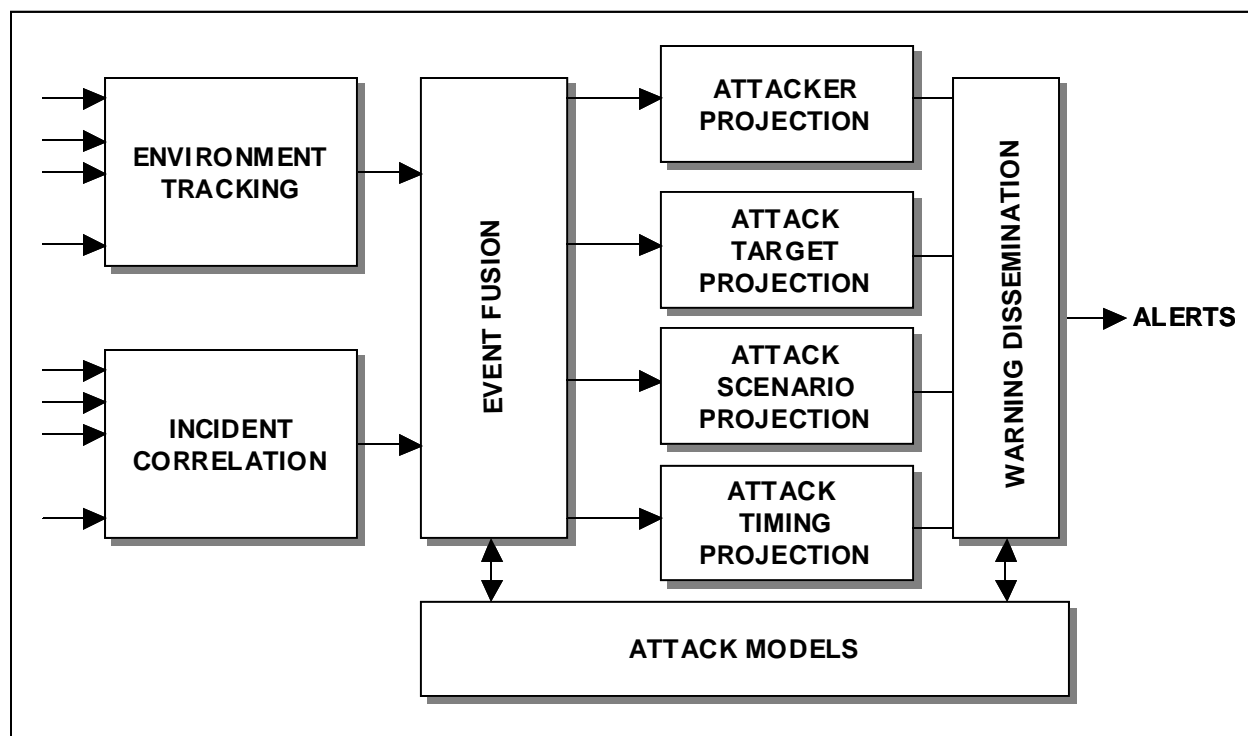
While this is undoubtedly a critical capability for maintaining an effective defensive posture, it is also the least mature. Discussion in the community seem to focus on the identification of precursors to attacks as “observables,” tracking a broad range of social, political, organizational, intelligence and technical events that can be fused with incident reporting to postulate attacker actions including attack target sites and systems, attack scenarios and timing. There is also a bent toward the use of various attack models to serve as a foundation for these projections.



**Figure 8.2-5 Possible Allocations of Detect and Respond Functions**

## 8.2.4.3 Functions to Support Incident Detection

Detection of incidents (or intrusions) is typical of a local site operation, as discussed in detail in Sections 6.4 (Network Monitoring within Enclave Boundaries and External Connections), and 7.2 (Host-Based Detect and Respond Capabilities within Computing Environment) of the Framework. In a broad sense, these functions at the local level are performed to determine the security posture and status of a local site (or environment) typically using network-based and host-based sensor technologies, supported by local analysts to identify vulnerabilities, intrusions, and malicious code attacks. Typical functions associated with this are shown in Figure 8.2-7.



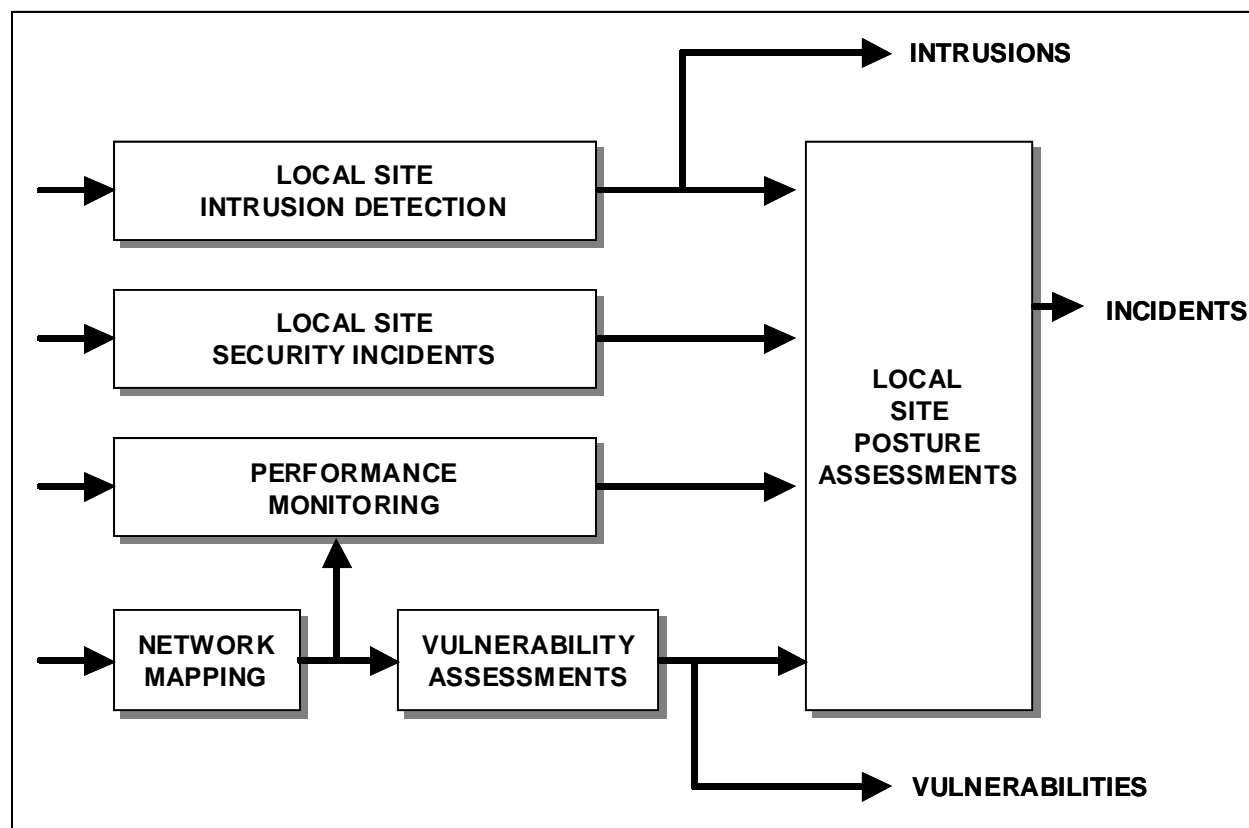
**Figure 8.2-6 Functions to Support Warning**

To be consistent with other functional structures discussed in this section, we distinguish this from incident characterization, where operators perform analyses to discriminate between alarms, events, interesting events, and intrusions. As inferred by the diagram, these functions go well beyond intrusion detection to consider security incidents, performance irregularities, and vulnerabilities identified by scanners or penetration (e.g., RED Team) testing.

## 8.2.4.4 Functions to Support Incident Characterization

These functions draw from the results of the incident detection discussed above to interpret the true nature and criticality of each alarm that is created by the local sensors. Typical functions with this are shown in Figure 8.2-8.

In addition to the primary inputs from incident detection, warning alerts provide an additional focus on specific attack sources and/or types of attacks. Ideally, the outputs of these functions would provide some sense of an intruder's intent, scenario, and the identification of the source of each incident identified. Typically, the results of these functions are used as input to the incident response functions, discussed below.

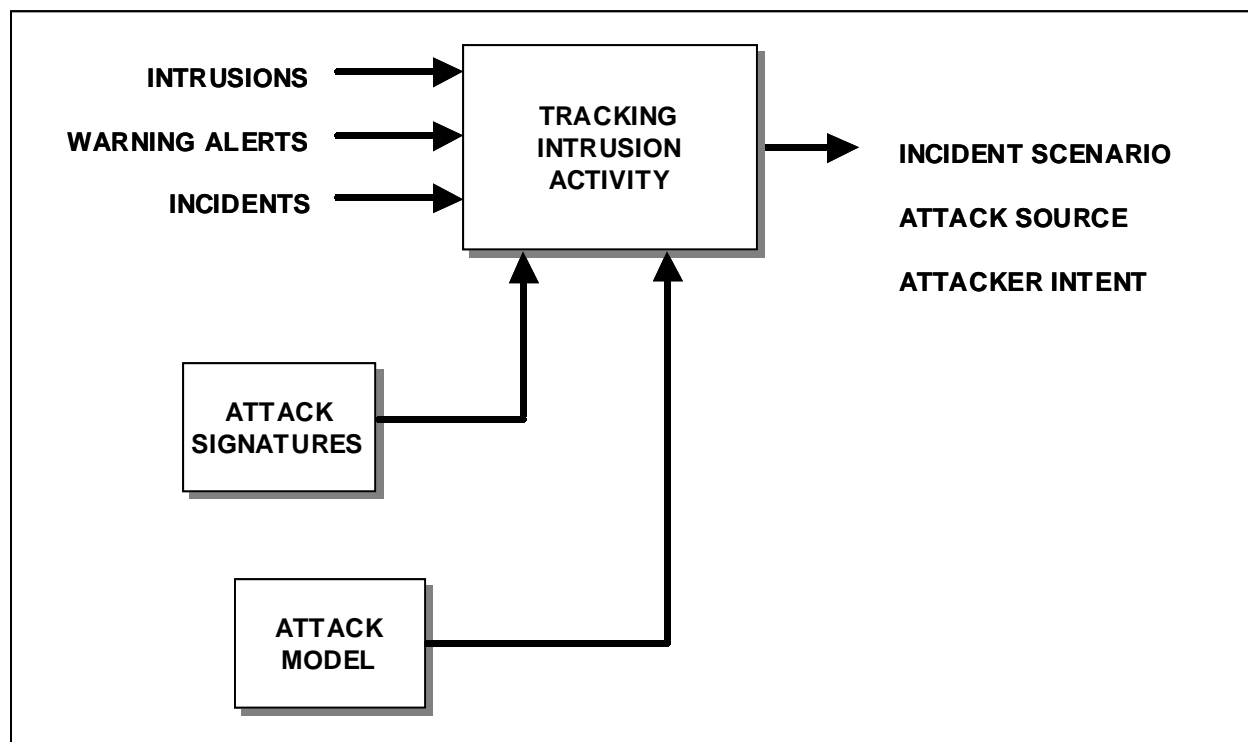
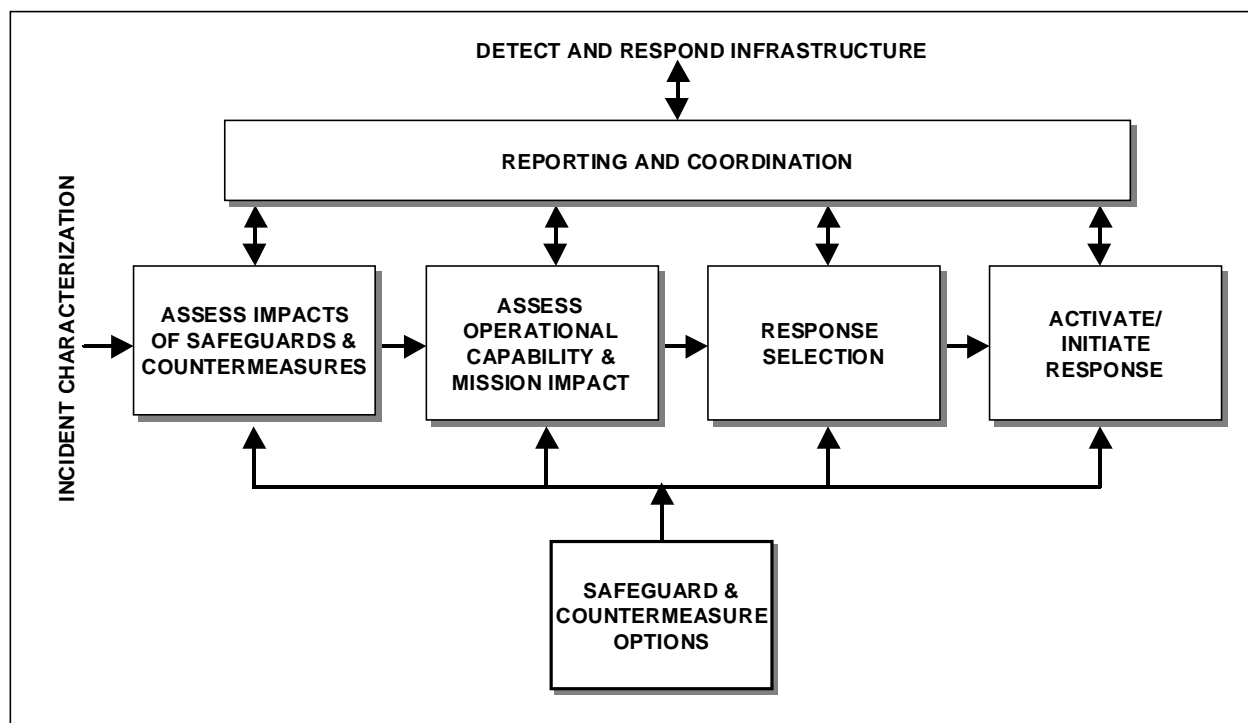


**Figure 8.2-7 Functions to Support Local Incident Detection**

### 8.2.4.5 Functions to Support Incident Response

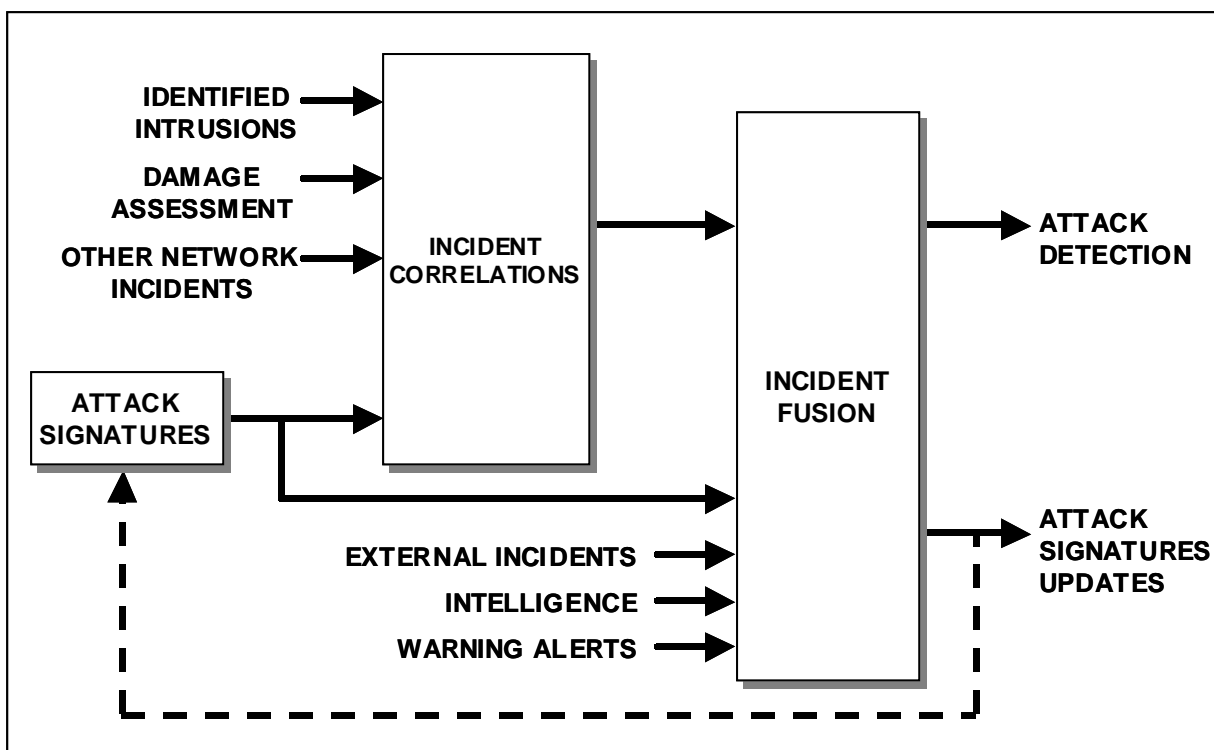
As discussed earlier, the local environment is ultimately responsible for executing a response to mitigate the effects of the intrusion and restoration to return the systems and networks to normal operation. Typical functions associated with this are shown in Figure 8.2-9.

These functions draw from a set of pre-established safeguards and countermeasure options. Selection of an appropriate response option would be made based on a number of assessments. These assessments first address the impact (and any anticipated progressions) of the incident on the site's operational capabilities and its ability to perform its missions. The focus is then turned to how the activation of available responses themselves would impact the site's operational capabilities and ability to perform its missions. Coordination with the Detect and Respond infrastructure (when appropriate) can provide recommendations about the technical impacts that response options may have on incidents associated with ongoing attacks as another factor for consideration in selecting a response. Finally, these functions include the activation of the selected response, typically intended to contain, assess damage, eradicate, reconstitute, and recover from the effects of the incident (or attack) to the local site capabilities.

**Figure 8.2-8 Functions to Support Incident Characterization****Figure 8.2-9 Functions to Support Incident Response**

## 8.2.4.6 Functions to Support Attack Determination

Building on intrusion and incident reporting from local sites and external events, these functions focus on determining if an attack (versus an incident) is underway or has occurred. Typical functions associated with this are shown in Figure 8.2-10.

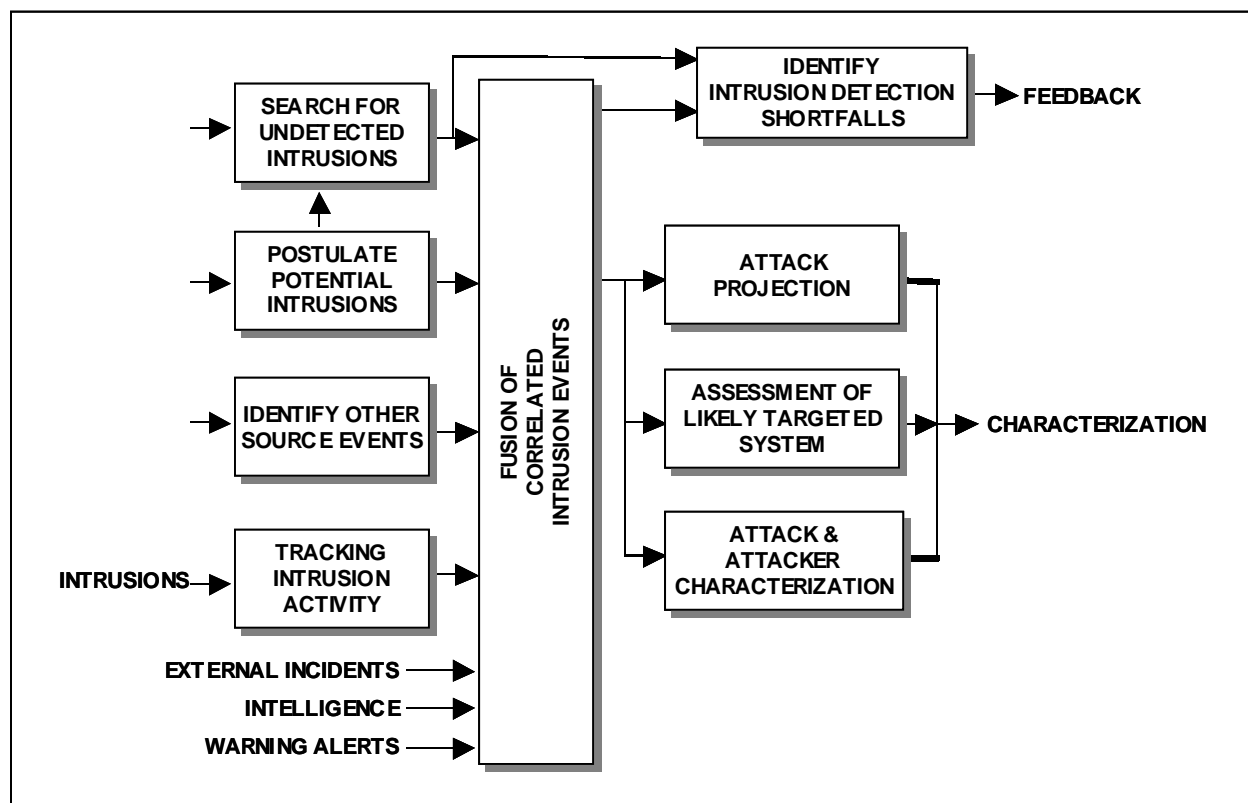


**Figure 8.2-10 Functions to Support Attack Determination**

Drawing from the local sensing functions discussed in Sections 6.4 (Network Monitoring within Enclave Boundaries and External Connections), 6.5 (Network Scanners within Enclave Boundaries), and 7.2 (Host-Based Detect and Respond Capabilities within Computing Environment) of the Framework, this activity also includes correlation of incident data from all sites within its constituency and fusion of that with warning alerts, all-source intelligence reports, and other external events to discern if an attack is underway.

## 8.2.4.7 Functions to Support Attack Characterization

Once the determination has been made that an attack has been detected, this set of functions focuses on analyzing the attack in term of its intent, approach, projections of how it will proceed, likely impacts, and possible identification of the attack source. Typical functions associated with this are shown in Figure 8.2-11. The functions can be considered in two categories. The first is fusion of the various sources of information to identify all relevant events and data to be analyzed. The second is a series of specific analysis functions that focus on the various aspects of the characterization.

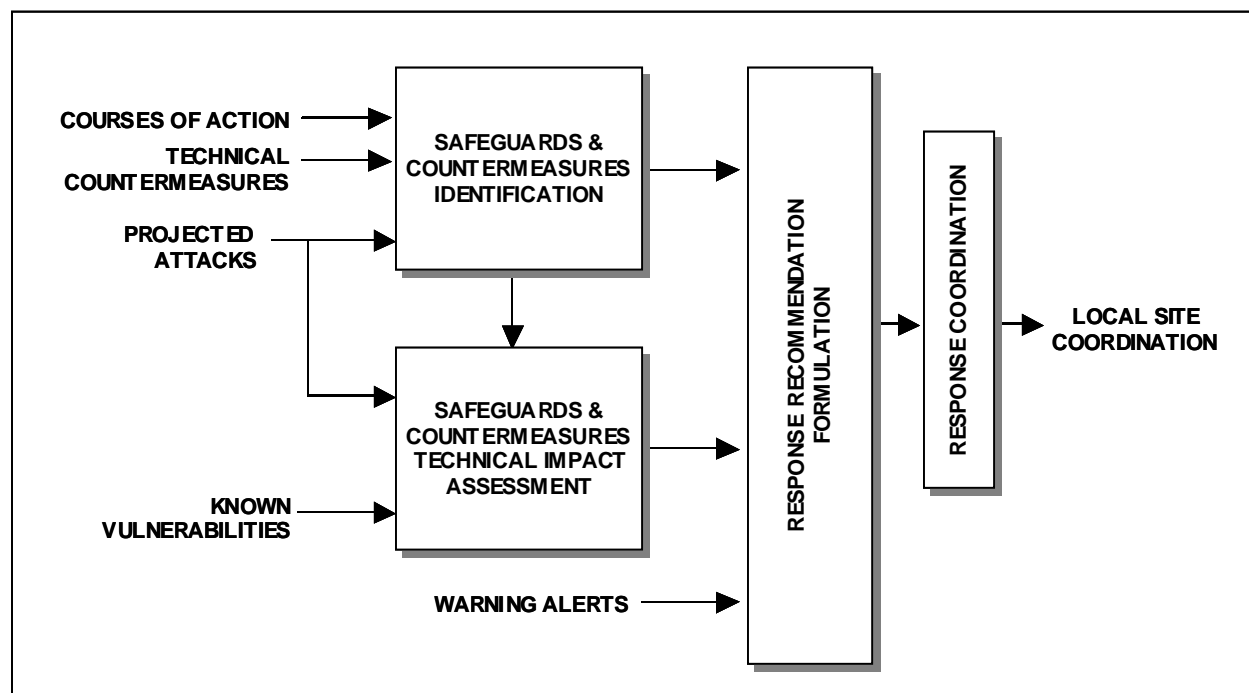


**Figure 8.2-11 Functions to Support Attack Characterization**

Resources available to support analysis include Warning alerts, all-source intelligence, external incidents, known attack scenarios, and attacker signatures and electronic fingerprints. A side benefit of these analyses is feedback that can be provided to local intrusion detection systems to support their tuning, updating their attack scripts, and the like, to improve their detection capabilities as they pertain to the ongoing attack.

## 8.2.4.8 Functions to Support Response Coordination

Once an attack has been detected and characterized, the real value the system infrastructure can provide is the coordination of an effective response at the local sites that will mitigate the effects of the attack and support the restoration needed to return the systems and networks to normal operation. Typical functions associated with this are shown in Figure 8.2-12. The thrust of these functions is to assess, on a technical (versus operational and mission impact) basis, the effectiveness of available pre-planned courses of action, safeguards, and countermeasures against the identified and projected attack scenarios.

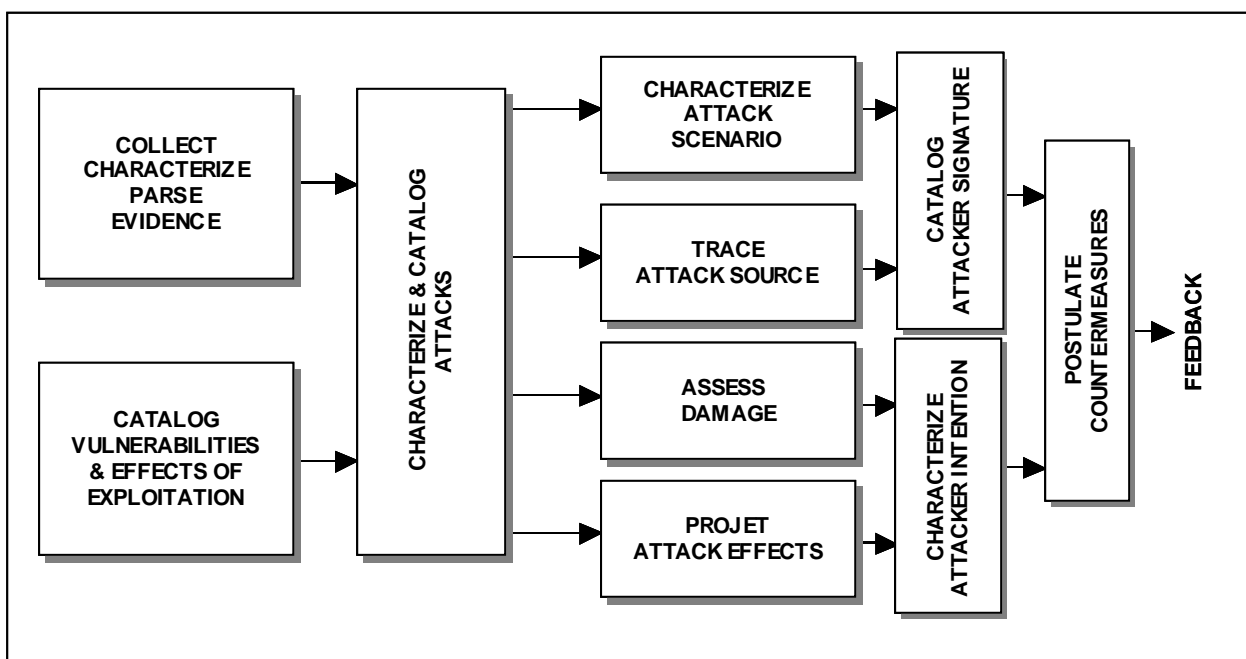


**Figure 8.2-12 Functions to Support Response Coordination**

Typically, local organizations and sites are in the best position to assess operational and mission impacts, based on projections of technical impacts to network services and system operations. Recommendations are formulated to assist local sites in the containment, damage assessment, eradication, and restoration to normal operational state. When appropriate, this also includes development or refinement of React mechanisms tailored to unique aspects of the ongoing attack.

## 8.2.4.9 Functions to Support Attack Investigation

This remaining set of functions focuses on analyzing how an attack was accomplished in order to provide feedback to improve existing (and future) protect, detect, and respond capabilities, ensuring that similar exploitations cannot occur. When appropriate, they are also structured to provide evidence when prosecution of attackers is pursued. Typical functions associated with this are shown in Figure 8.2-13. These functions are typically performed post attack with extended timeframes available for in-depth analyses. They can be considered in four basic groups or categories. The first is to establish and maintain a catalog of known vulnerabilities and the effects of known exploitations that provide a foundation for those analyses. These can include determining the effects of known attack sequences and potential modifications to those attack sequences. The second group, which is the primary focus for attack investigation, addresses characterization of the attack and attacker built from any available cyber evidence (e.g., audit logs, Transport Control Protocol (TCP) dumps).



**Figure 8.2-13 Functions to Support Attack Investigation**

When required, this also provides evidence that could be used in subsequent prosecutions of attackers. The third establishes a set of attacker “signatures” (which could be thought of as a fingerprint file) that can be referenced when investigating future attacks. The remaining group focuses on developing and providing feedback for improving countermeasures and safeguards.

## 8.2.5 Relevant Detect and Respond Technologies

Cyber attack detection and response technologies (predominantly focused on intrusions) have emerged within the last several years due in large part to situations that stem from the worldwide interconnectivity created by the Internet. The computer literate can gain access into government and commercial internal networks via public routes using software hacking tools that can be easily downloaded from the Internet.

The previous section provided a perspective on the types of functions that are typical for various layers of a Detect and Respond infrastructure. This section provides guidance on technologies that are available to implement these functions and considerations for their selection and effective use. The section concludes with a reference model that provides an overall context for these technologies in a Detect and Respond infrastructure setting.

The defense in depth strategy and the overall Framework reinforce the close relationship of personnel, operations, and technology in realizing an effective Information Assurance (IA) posture. This cannot be emphasized too strongly across the Detect and Respond disciplines. When looking at the state of Detect and Respond technologies, it is clear that there are no “easy answers.” Many of these technologies really provide measurement (instrumentation) capabilities that must be interpreted by highly skilled analysts. Other technologies provide tools to support the analysis operations. Even the response technologies require well-trained



and highly skilled operators to ensure that the response mitigates, rather than exacerbates the effects of an incident or attack. Three major issues associated with effective technology deployment are:

- *Where* in the network they are deployed to ensure they address critical network resources?
- *How often* they are used based on the operational concepts of operation and availability of operators and analysts?
- *What skills* the operators and analysts must have to make effective use of the results?

It cannot be over-emphasized that unlike protect technologies, Detect and Respond technologies do not in themselves offer any real protection. Rather, they enable the processes and functions that can mitigate the effects of an attack and restore the information systems and networks to an operational condition.

### 8.2.5.1 Technology Categories

Although commercial intrusion detection products have been available for several years, a number of recent and highly publicized hacking cases have created a renewed interest in the broader field of Detect and Respond technologies. Research by government, industry, and universities is ongoing to determine what constitutes an attack, how to detect an attack, and how to respond to an attack.

Today, most technologies tailored for Detect and Respond use provide information to an analyst, assist an analysis, or provide a means for responding based on the results of the analysis. Figure 8.2-14 shows the broad range of technologies that are addressed in this section of the Framework.

### 8.2.5.2 Monitoring and Scanning Technologies

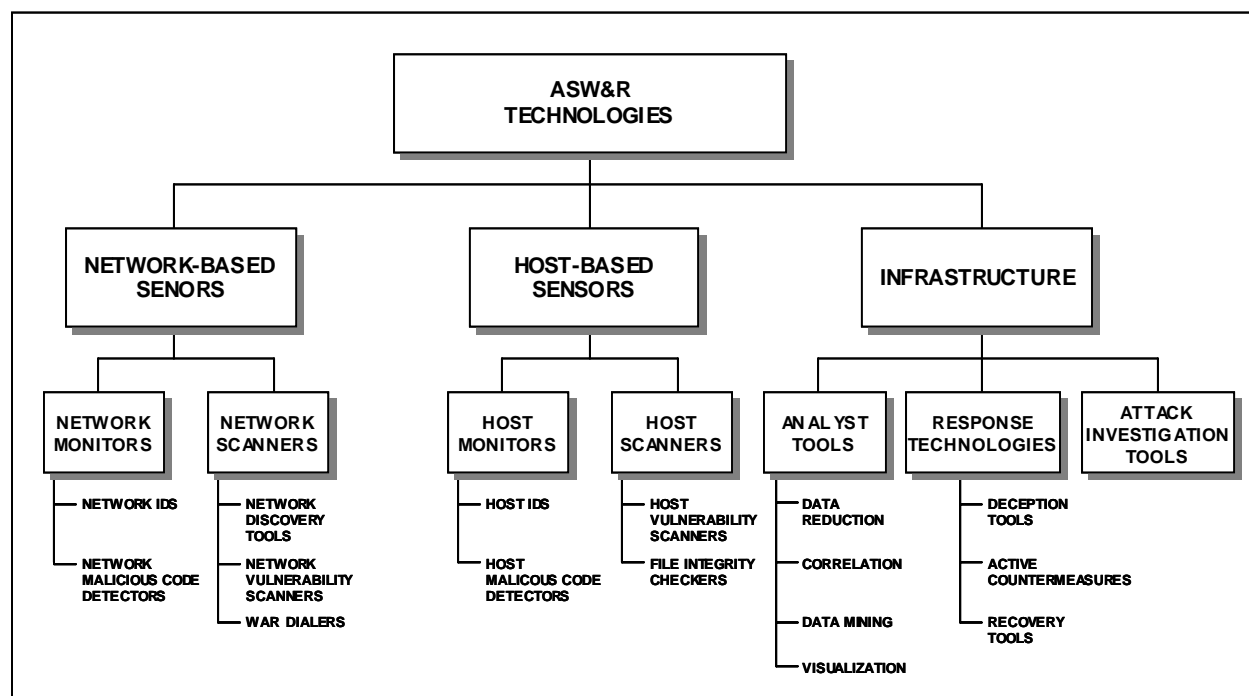
It should be noted that monitoring and scanning technologies (characterized broadly as sensors) are covered in depth in other sections of the Framework. Specifically Sections 6.4 (Network Monitoring within Enclave Boundaries and External Connections) and 6.5 (Network Scanners within Enclave Boundaries) address network-based monitoring and scanners, respectively, while Section 7.2 (Host-Based Detect and Respond Capabilities within Computing Environment) addresses host-based sensor technologies. This material is synopsized in this section to provide a context for the remaining technologies and to facilitate discussions of when and how to use these technologies in a synergistic fashion. Figure 8.2-15 identifies the general categories of these technologies.

## Technology Overview

Network and host-based sensors provide alerts and supporting information to network operators and administrators that a vulnerable condition exists or an event has occurred within the enterprise and thus creates an opportunity for them to analyze and evaluate when actually transpired. This allows an appropriate action (as specified by the security policy for the organization) to be initiated. If the attack is detected in real time, it may be possible to mitigate the damage resulting from the attack. If detected after the attack is over, the logging features of

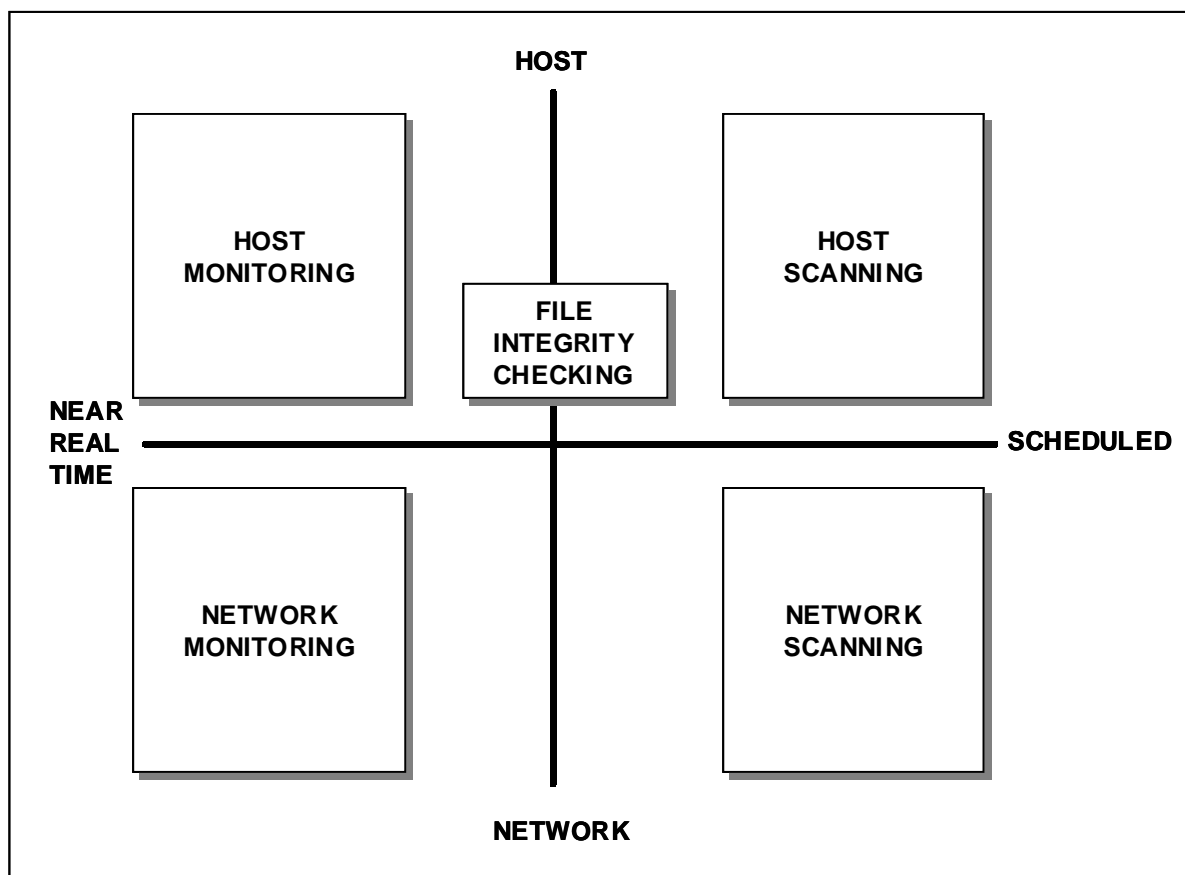
the sensors may identify why the attack was successful so that exploitable weaknesses can be fortified.

## Monitors



**Figure 8.2-14 Detect and Respond Technologies**

*Network IDSs* examine traffic on the wire in real time, examining packets looking for dangerous payloads or signs of abuse (e.g., malformed packets, incorrect source or destination addresses, and particular key words) to spot attacks before they reach their destination and do the damage. When suspicious activity is identified, a network-based IDS is capable of both raising alerts and terminating the offending connection. Some will also integrate with the firewall, automatically defining new rules to shut out the attacker in the future. As indicated in the earlier sections of the Framework, the high incidents of false positive detection make automated response mechanisms undesirable. Network-based IDSs typically operate on independent computers so there is no impact on the performance of mission systems. They are typically deployed one per network segment, since they are unable to see across switches and routers.



**Figure 8.2-15 Sensor Technologies Grouping**

*Host Intrusion Detection* provides an agent that resides on each host to be monitored. The agent collects information reflecting the activity that occurs on a particular system. The monitor scans event logs, critical system files, and other auditable resources looking for unauthorized changes or suspicious patterns of activity. When anything out of the ordinary is noticed, alerts or SNMP traps can be initiated automatically. The agent may also behave in a similar manner to the network-based IDS in that it will examine packets on the wire to compare against a database of known attacks—but in this case it is restricted solely to packets targeted at the host machine. For this reason, it is ideal in a highly switched environment, to protect specific critical servers, or for otherwise heavily loaded networks (where it may be difficult to protect the entire network. Some host-based IDSs also include a “personal firewall” capability to provide additional protection for the host machine. Unlike its network counterpart, Host IDSs operate on mission critical systems, and as such do impact their performance for mission operations.

*Malicious Code Detectors* prevent and/or remove most types of malicious code. The use of malicious code scanning products with current virus definitions is crucial in preventing and/or detecting attacks by all types of malicious code. Malicious code detectors should be implemented across the enterprise. Defense against malicious code is only as good as its weakest link; if one system can be compromised, then the entire enterprise is at risk. Centralized management for the AV capabilities with a common set of policies is strongly recommended.

## Vulnerability Scanners

The Framework makes the distinction between scanners and the monitoring devices discussed above. Monitors typically operate in near real time and tend to measure the effectiveness of the network's protection services in practice since they are subjected to actual exploitation attempts. Scanners, on the other hand, are preventative measures, typically operating periodically (or on demand) to examine systems for vulnerabilities that an adversary could exploit, evaluating effectiveness of the system infrastructure's protection. Vulnerability Scanners sometimes referred to as "risk assessment products" provide a number of known attacks with which network administrators can probe their network resources proactively. Scanners perform rigorous examinations of systems in order to locate known problems that represent security vulnerabilities.

*Host-based Scanners* use an agent loaded on a system to examine a server or client. This examination can determine the potential system-level vulnerabilities that exist on a particular system based on known vulnerabilities in the operating systems. These technologies typically connect into a management console that can report on the status of all systems with agents across the network.

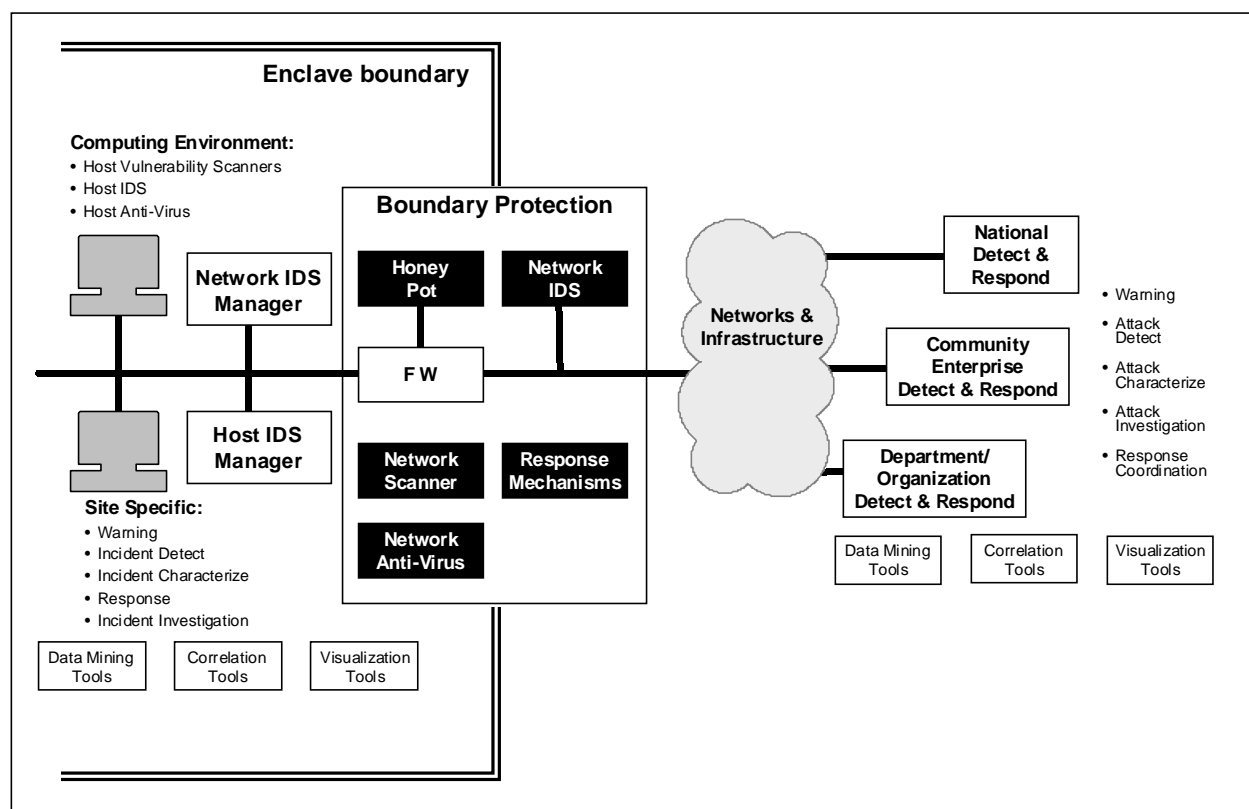
*Network-based Scanners* examine a network and take inventory of all devices and components within the network infrastructure. These components, the network configuration, and the various versions of software controlling the network are examined and compared to a database of known vulnerabilities.

*War Dialers* are a specialized type of network vulnerability scanner technology. Once identified, those backdoors can be closed or some type of security plan created to preclude use of that particular point of entry. Along with a strong modem policy describing the need for modem registration and PBX controls, War Dialer scanning can help an organization defend itself against such dangers. Use of this type of technology can help an enterprise to identify vulnerable backdoors (e.g., unsecured modems across an enterprise) before an attack occurs.

*File (Software) Integrity Checkers* are a specialized type of host scanner technology that verifies the integrity of files detecting when files have been changed. As with the host vulnerability scanner technologies *discussed* above, these technologies tend to run off-line, and thus are not a protection mechanism. Typically they operate periodically, based on an event (e.g., file access) or on demand.

## Considerations for their Deployment and Operation

Deploying combinations of both network and host-based sensors provides the best possible security by monitoring network-based traffic and host specific exploitations directly on target workstations. This combination provides significant attack protection and facilitates policy enforcement for any size enterprise. Figure 8.2-16 identifies potential locations for their deployment.



**Figure 8.2-16 Possible Sensor Deployment Locations**

When possible, it is recommended that the sensors be linked into the overall system and network management capabilities for an enterprise-wide solution. This eases individual sensor management, facilitates central reporting, and provides a more coherent perspective on the status of the enterprise overall.

Malicious code detectors should be implemented across the enterprise, on every system and network. Most of these technologies provide a means for sending responses or alerts at the server level, and some at the console level. It is always desirable to notify anyone that may have been infected that malicious code has been detected.

If scanners are deployed, it is important to consider what and when scans are performed. Otherwise, it is possible that mission critical servers become busy responding to simulated attacks during times of peak demand. Assessment frequency is a factor of how often network changes are made as well as the security policy for the enterprise.

The most important aspect to consider for integrity checker operation is deployment timing. To be their most effective, integrity checkers should be initialized on systems before they are placed into production and made generally accessible to their user communities. If they baseline monitored files and data structures anytime after a system has “gone live”, it is possible that the system has already become compromised and the integrity checker will miss changes that have already occurred.

### 8.2.5.3 Analyst Tools

Many intrusion detection and vulnerability scanning tools described above and in previous sections of the Framework come with their own rudimentary analysis tools. Some third party vendors offer tools that will input security audit logs and intrusion event logs from some systems for further analysis, particularly if they have been generated in some standard format (e.g., Open Database Connectivity). The interoperability standards for some of these formats (Intrusion Detection in particular) are still under development in standards bodies and Government-sponsored activities such as DARPA's Common Intrusion Detection Framework (CIDF) program, the Internet Engineering Task Force's (IETF) Intrusion Detection System Working Group and the ISO SC27 standards group.

### Technology Overview

While network and host sensor technologies have been developed specifically for Detect and Respond functionality, analyst tools have evolved from more general-purpose applications. While basic tools and technologies exist, commercial analyst tools have not in general been tailored to this environment. We note that the Government sector (e.g., the Intelligence Community) has developed a number of custom tools that more closely relate to this use, however they are considered beyond the scope of the discussion in this Framework.

In order to support the analyst in performing the functions described in Section 8.2.4 Detect and Respond Functions, tools and techniques must be assembled that allow analysts to use all aspects of the information analysis technologies discussed below across the problem. The kind of tools required to do the "all source" type of analysis required by the Detect and Respond infrastructure are not currently available in the commercial sector, but any analyst tools (individually or in combination) must provide functions in the following areas:

*Data Reduction* – IDSs are notorious for generating large amounts of mostly superfluous information if not configured precisely. Even when well configured, their design is such that the system errs on the side of identifying, tagging, and reporting on *all* potential intrusion events. This data must be reduced down to information of import before any additional analysis steps can be performed. Often, data reduction takes place incrementally during many of the analyst functions described in Section 8.2.4 (Detect and Respond Functions). Models of "acceptable behavior" are typically used to reduce information. Local knowledge, such as configuration of the networking environment, knowledge of the application and systems in use across the network or enclave, and the expected traffic patterns of normal behavior, can all be used to reduce the mass of information generated by these systems to more manageable and germane levels.

*Data Correlation* – Correlation of events over a large set of data, even after data reduction techniques have been applied, to identify problems or determine if attacks are underway can be time consuming and place extreme demands even on experienced operations staff. The larger the correlation environment, the more complex and detailed such correlations become. Often, operations staff cannot keep up with the increasing rates at which events are generated. Therefore, automated event management and correlation systems that can scale to large and complex environments are needed to accurately model and store the diagnostic knowledge possessed by operations staff. They must provide algorithms that analyze this knowledge in the context of the current system state to detect problems as they occur. Such systems must be

able to input and correlate data from disparate sources, from intrusion detection event data to external alerts and intelligence databases. Generally, automated correlation tools determine relationships among data by implementing one or more of the following reasoning techniques: Rule-Based Reasoning (RBR), Model-Based Reasoning (MBR), State Transition Graphs (STG), Codebooks, and Case-Based Reasoning (CBR).

RBR techniques may not be well suited to larger, enterprise-wide environments, but can work well in small domains, perhaps on the local level. Codebook reasoning is faster than rule-based reasoning given its streamlined encoding methodology, and is better suited for larger enterprise environments. STG techniques are limited to correlated events in a single object, and can't determine when problems occur across related objects. MBR also doesn't function well in large domains, and CBR does not scale well due to the need for a general case library, which would be different for each enterprise/local environment. A scaled approach based upon these techniques has yet to be developed.

*Data Mining* – Data Mining refers to capabilities to drill down through a database and display information in a meaningful way. It is one segment of the broader knowledge discovery technology that addresses knowledge creation overall. Data mining technology and techniques can be applied to the analysis environment with the goal of turning information from all sources in the Detect and Respond infrastructure into the identification of hidden attacks, patterns of attacks, and prediction of attacks. Data mining technologies can potentially discover hidden predictive information in large data sets. They use knowledge discovery, pattern recognition, statistical data analysis, and database systems technology to automate the search for information in data sets. Data mining technologies collect and analyze information from multiple data sets and check them for data integrity. They provide a clearer resolution of the information, provide an understanding of attacks in progress, and predict patterns of attacks.

Some specific work is already underway at Columbia University, where researchers have defined and tested a data-mining framework for adaptively building intrusion detection models. Their work uses auditing programs to extract information to detail each network connection or host session. Then they apply data mining techniques, such as classification, meta-learning, association rules, and frequent episodes to learn rules that accurately capture the behavior of intrusions and normal activities. These rules can be used build new detection models. While this is only part of the solution, it illustrates how data mining techniques are becoming an integral aspect of more advanced Detect and Respond tools base.

*Visualization* – Data visualization cuts across all the aforementioned areas. Technologies must be employed that make use of simple, yet effective visualization techniques to assist the analyst through the various functions associated with the framework. The use of common metaphors and design elements provide the ability to visually process presented information effortlessly. Gestalt principles of proximity, continuity, similarity, symmetry or good form, and closure, as well as the introduction of appropriate perspective and relevant color, all significantly enhance the analysis functions.

## Considerations for Their Selection, Deployment, and Operation

All the above factors must come together in a tool or series of technologies that provide to the analyst the ability to support the Detect and Respond infrastructure as described in 8.2.4 (Detect and Respond Functions). Numerous tools exist that provide partial solutions, but there are still

many challenges relating to common data export formats, the development of accepted reference models, and the problem of all-source data fusion, that allow a focus on attacks versus incidents.

These technologies become of critical importance in the context of an overall enterprise management strategy, particularly as it pertains to Detect and Respond operations. Today, many event management functions are handled manually. Analysts and operators monitor and correlate events and handle identified problems (or potential problems). This manual processing does not scale to the growing speed, complexity, and size of many enterprise networks. Utilizing these technologies, an enterprise management capability can accurately model and store diagnostic knowledge possessed by operations staff, and provide algorithms that utilize this knowledge in the context of the current system state to monitor, detect, characterize, and react to events in an efficient and effective manner.

Essentially, all these technologies must:

- Operate on a common data/information format. Given the nature of the tools required and the information to be processed, some sort of data warehouse construct is probably the most viable approach.
- Provide different levels of functionality at different tiers of the framework. Some tool functionality, such as the requirement to integrate event information with intelligence data, won't be required at a local level, but will be necessary at the organizational and national levels, particular where coordinated attack determination analysis is underway.
- Provide seamless operator interfaces between technologies and a common, yet flexible, visualization approach

There are no commercially available tools that provide all the necessary functions to satisfy the analysis needs within the Detect and Respond infrastructure. While there are fusion tools that have been developed within the Government that provide functions similar to those needed for the Detect and Respond environment, they do not synergistically bring together various analysis technologies in a single packaging for this specific focused purpose. For the most part, they have evolved and have been tailored for specific community (e.g., warfare and intelligence) operations. In some cases, there are efforts underway to adapt them to the Detect and Respond environment, however they have not reached the state of commercial technology offerings. Simple COTS approaches will undoubtedly require tailoring and integration efforts to build a cohesive shell or framework system around the various critical technologies.

## 8.2.5.4 Response Tools

There are two general classes of response tools considered within this Framework. One is a Deception Server, as discussed below. The second class of response tools, referred to as Active Countermeasures, focuses on implementing immediate mitigation actions to repel or redirect active attacks in order to minimize damage or re-establish and recover blocked or disabled services.



## Deception Servers

These response tools provide capabilities for characterizing and refining information pertaining to attacks in progress or particular attackers either by re-directing or luring attackers into highly instrumented system infrastructures designed to closely audit all activities. These systems are typically called Deception Servers, although they are more commonly known as honey pots, fishbowls, and upon occasion, Venus flytraps.

### Technology Overview

The concept behind deception servers is to present a “false” front, an instrumented server environment, with simulated well-know vulnerabilities (the “honey pot” construct) to lure attackers in with the promise of an easy score. These systems are designed and configured to emulate a production environment, but are in reality set up to alert network administration and security staff while at the same time generating detailed activity logs of the attack or intrusion event. The system thoroughly measures and tracks the would-be intruder’s activities.

While not a new idea, this *is* a relatively new class of product to be offered commercially. These products are capable of simulating a range of different network servers and devices in order to act as an attractive decoy for the would-be attacker. While the attacker concentrates on the decoy services, the honey pot collects as much evidence as it can while it is alerting the administrator.

When an incident is detected it is the organization’s choice to terminate the connection immediately or to continue to allow the attacker to explore the façade system. If the connection is terminated, the attackers knows that they have been detected and may try a different approach or to attack a different organization with the same attack. If allowed to continue unchallenged within the deception environment, information about the attacker can be gained. This information can be recorded and used by law enforcement officials to apprehend the attacker and take suitable legal action.

Deception servers can be useful only if the environment being protected has sufficient resources to use them once they are deployed.

### Considerations for Selection Unique to the Deception Server Environment

Besides the usual criteria for selection of any software package or technology to be used within the Framework, such as supportability, dependability, clarity of user interface and documentation, ease of use and the like, there are a few fairly unique aspects to consider. There are a number of considerations that should be taken into account when choosing a deception server product for deployment

***Platform and Emulation Operating System*** – The most important factor to consider is platform support. The system should either run on the same type of platform that is commonly used in the environment it will be protecting, or emulate the operating system that is running on the true production systems that surrounds it. Depending upon the target environment, Windows NT and various versions of Unix should be supported. Some products will even attempt to emulate network appliance services such as Cisco Internet Operating System (IOS).

***Commercial Product vs. “Home Grown”*** – There are numerous documents available in the community that describe how to configure a deception server from base operating system

installations. This could be considered as a cost savings option, particularly if there are operating system support personnel available. However, it may be much more efficient to simply use one of the available products “out of the box.”

**Emulation Level** – Some deception servers attempt to emulate more commonly offered network services while others emulate the application level. The closer the emulation to the true implementation, the more likely the ruse will work without alerting the attacker to the deception. One available technology actually makes a copy of your production system environment, securing it and instrumenting it on a second hardware platform for deployment as a deception server. Those systems that only emulate at an application level are susceptible to network-level Operating System identification tools, such as the commonly used Nmap. The level of deception required depends upon how high the risk factors are for the environment and the probability of threats coming from highly sophisticated attackers. For environments with few resources, easily deployed, commercially available emulation packages should suffice. However, for the best coverage, a full-blown dedicated system that imitates the production environment in every way will provide the best protection possible.

**Reporting and Logging** – Of course, the depth and breadth of logging are important, particularly based upon what the true operational goals of the deception server are. If the goal is to simply be alerted to the fact that an intrusion is underway and provide some level of data to assist in the foiling of the intrusion and recovery, the level of audit and reporting need not be particularly high. However, if the goal is to provide sufficient evidence to law enforcement officials to trace and potentially prosecute an attacker, a higher level of audit, reporting, and supporting documentation are required.

## Considerations for Deployment and Operation

There are a number of considerations for deployment and operation of deception servers.

**Placement on the Network and Re-direction** – Several methods exist for placing deception servers into a network infrastructure and ensuring attackers go after it. For example, one can either set up boundary routers or firewalls to re-direct non-production services (e.g., FTP or Telnet) to the deception servers rather than to just not support them, and then route normal services, such as HTTP, to production systems. The drawback, of course, is that if attacks take place using production services, the deception server provides no added value. Another approach is to place a deception server at the same logical network level as production servers and have it emulate full production services, so it can become targeting in attack “sweeps”.

**Legal Issues** – Little or no legal precedence has been established for deception servers. If deception servers are deployed, some potential liabilities could be experienced. It would be wise to post the same restricted use notifications that are found on the enterprise’s true production systems. Additionally, be prepared that if the deception server is compromised and then subsequently *used* as a stepping off point for attacks elsewhere, the organization that deployed the deception server could be found culpable, more so than if their normal production servers were compromised despite due diligence efforts. It should be kept in mind that deception servers are detection tools and be treated as such, and unless the deploying organization is a law enforcement agency, unfair entrapment charges can’t really be made successfully.

## Active Countermeasures and Recovery Tools

Active countermeasures and recovery tools focus on terminating the intrusion or attack and restoring affected services or lost data as soon as possible. Recovery may also include initial (technical) damage assessment tools that ascertain the extent of the damage inflicted during the intrusion or attack. These should be differentiated from attack investigation tools, which are used to gather information about intrusions with the intent, among other activities, to trace, locate, apprehend, and prosecute intruders and attackers (addressed in subsequent sections).

### Technology Overview

***Reconfiguration, Containment, and Disconnection Technologies*** -- There are numerous approaches to initiating active countermeasures that serve to halt or block attacks that are discovered against an environment. Typically, there are no tools one can acquire that stand-alone and are used to repel attacks. Most countermeasures come bundled with intrusion detection systems. They provide either a standalone capability (e.g., the ability to send TCP disconnects to certain active connections determined to be the source of attacks), have programmed interfaces to network equipment (switches, hubs, routers and firewalls) so certain connections can be cleared or blocked at the network level, or allow new filtering rules to be instituted based upon addressing or protocols associated with the attack. Many tools allow the creation of pre-canned scripts that can be executed causing dynamic reconfigurations across the enterprise.

Additionally, some host-based tools provide the ability to interface with the host operating system to allow quick disabling of accounts that are being used as launch points for attacks. Dynamic access control modifications are also possible. All these tools should be focused upon minimizing the period in which the attack takes place, and consequently minimizing the damage, either from the original attack or as the intruders attempt to cover their tracks as they back out.

These tools (or more appropriately features of available intrusion detection tools) must be chosen carefully and their use within the secure infrastructure planned accordingly. Each of the various attack mitigation features should be thoroughly tested to ensure that they do not wreak more havoc on the enterprise than the original attack. Some tools allow the automatic institution of countermeasures. It is recommended that automatic “shunning” not be implemented until all scenarios are tested and sufficient operational experience in the particular environment indicates the risks are minimal.

***Recovery Tools*** – Damage assessment and recovery tools include disk repair and recovery tools as well as operating system specific tools that are able to make repairs to OS-specific data structures on the system (e.g., the Windows registry). It is important to prepare these tools ahead of time, in anticipation of having to recover from attacks, since no protection features are foolproof.

Backup recovery tools are an important component of this part of the framework. Each set of tools must be chosen to work with the particular platforms and information system applications running within the enterprise. Pre-event planning and rehearsals should be conducted to ensure that the tools are configured appropriately and operations personnel are sufficiently trained. Processes and procedures for proper backup execution, testing, and the selection of the appropriate periodicity to execute backups are all critical factors in the pre-

planning phases of recovery operations. Some of the file integrity checking tools addressed in Section 7.2.4 (Host Scanners – File Integrity Checkers) can also be used in the recovery process, determining which files may have been corrupted during the attack and may have to be restored from protected media. Besides the technology, appropriate planning is an absolute necessity as part of any response capability.

## 8.2.5.5 Attack Investigation Tools

Also referred to as computer forensics tools, attack investigation tools, and computer forensics science in general, focuses upon acquiring, preserving, retrieving and presenting information associated with illegal intrusion activities. Three roles of a computer within a criminal context have been identified. The first is where the computer is a target of an attack or intrusion. The second is where a computer is used as an instrument of an attack (a hacker's computer, for instance). The third is where a computer may be a repository for information pertaining to the commission of a crime, containing databases, images, etc.

In the context of the Detect and Respond infrastructure, attack investigation or forensics tools consider the first and second roles. The first, where the computer is the subject of the attack, and the second, where a third party computer was attacked and usurped, then being used in subsequent attacks upon other systems. The aspect of seized computers being examined for their role in criminal activities, whether as a tool or as a repository, is beyond the scope of this section of the Framework.

## Technology Overview

There are three general phases to any computer forensics process: acquisition, examination, and utilization, and consequently different tools for each. In the acquisition phase, information must be acquired from the systems that have been intruded upon and/or attacked in such a way that all the information on the system is captured. In situations where criminal prosecutions are a goal of the investigation, the information must be collected and maintained consistent with rules of evidence. In the examination phase, appropriate tools must be used to analyze the information on the system with the intent of attempt to ascertain such facts as:

- How the attack was achieved (i.e., what vulnerability, technical or procedural, was exploited),
- What information the intruder may have left behind to implicate themselves (e.g., trace logs, malicious code or Trojan Horse software, trademark methods, system damage),
- What the intent of the intruder was (e.g., exploration/curiosity, malicious damage, information theft, denial of service, service theft).

Finally, the utilization phase of the forensics process allows for the creation of formal reports, the certification of the chain of custody thread, and all other aspects that then allow the pursuit of a criminal investigation leading to a potential prosecution.

Most standard computer forensics tools focus on the preservation of evidence, the analysis of information for criminal activity, and then the final packaging for prosecution. In the Detect and Respond infrastructure, while many of these tools have applicability, additional analysis tools that focus upon log and event analysis are also important, particular those logs and events

from secondary systems such as routers and firewalls, and not necessarily just the pilfered target system itself. However, in all cases, rules of evidence must be followed to support successful prosecutions.

When a situation arises in the Detect and Respond environment where attack analysis is intended to potentially lead to criminal prosecution, acquisition tools that capture and preserve the evidentiary trail of information must be used instead of simple log or event information capture and copying. Tools that make exact, certifiable copies of information, often entire disk images, must be deployed.

For analysis, tools that not only attempt to recover lost or deleted information (an intruder covering his/her “tracks”) must be deployed, but tools that analyze log events and audit information to build a profile of how an intrusion progressed must also be applied. If necessary, tools that can analyze down to individual Transport Control Protocol/Internet Protocol (TCP/IP) segments and datagrams (TCPdump) must be used along side the more traditional computer forensics tools.

Finally, tools that generate reports, document the chain of custody, and just generally provide additional efficiency, fill out the third phase, utilization.

## Considerations for Selection and Operations

There are a number of factors associated with attack analysis that should be considered when pulling together a stable of appropriate tools.

*Ease of Use and Integration* – A clean, robust user interface, particularly in the complicated analysis phases of an investigation, is critical. Many tools handle all aspects of attack investigation (acquisition, analysis and utilization) in complete packages, most focused upon computer crime scene investigation. It is important to consider if these all-in-one packages adapt easily to the operational environment in question. Also, in most cases, a long, drawn-out investigation will have prohibitive impact upon operations. The speed with which information can be collected for later analysis is critical.

*Preservation of Evidence* – Do the tools preserve evidence appropriately, per acceptable law enforcement or prosecutorial standards. Do the disk copying or information copying tools function in such a way as to ensure a perfect copy is preserved. What tradeoffs are available between speed and copy perfection (full image versus information copy).

*Flexibility* – Can the tools collect, preserve and analyses information from the systems deployed in the local environment.

*Operational Approach* – The available tools are still in focused mostly upon single activities, such as information capture or disk imaging, log analysis, the discovery of deleted files or hidden information, etc. Consequently, particularly in a Detect and Respond situation, and well-composed investigative framework must be established ahead of time, to provide the context for the implementation of the tools. The functions during an investigation are described in Section 8.2.4.9 Functions to Support Attack Investigation above, but the next level of detail appropriate to the particular environment in question, such as operations personnel availability, budget, local and/or national policies on how long systems can remain offline for investigation purposes, etc., all must drive the particular tool acquisitions.

## 8.2.5.6 Related Detect and Respond Operational Considerations

While there are a number of technologies available to support various aspects of Detect and Respond, there are also important considerations that deal with their selection, deployment, and operation. Some of these are discussed below.

### Independent Testing of Technologies

Another factor slowing the development of these technologies is the lack of adequate testing and product certification facilities. Large-scale testbeds are needed to test these systems using real-world simulations and to develop metrics, verification procedures, and standard test-case scenarios. There is a real need for independent laboratories to evaluate and certify products, providing unbiased and accurate evaluations of relevant technologies that can be made available to network security customers.

The National Security Telecommunications and Information Systems Security Policy (NSTISSP) No. 11 provides the national policy that governs the acquisition of IA and IA-enabled information technology products for national security telecommunications and information systems. This policy mandates that effective January, 2001 preference be given to products that are in compliance with one of the following.

- International Common Criteria for Information Security Technology Evaluation Mutual Recognition Arrangement.
- NSA/NIST National Information Assurance Partnership (NIAP).
- NIST Federal Information Processing Standard (FIPS) validation program.

After January 2002, this requirement is mandated. Department of Defense (DoD) Chief Information Officer (CIO) Guidance and Policy Memorandum No. 6-8510, Guidance and Policy for Department of Defense Global Information Grid Information Assurance reference this same NSTISSP No. 11 as an acquisition policy for the Department.

The International Common Criteria and NIAP initiatives base product evaluations against Common Criteria Protection Profiles. NSA and NIST are working to develop a comprehensive set of protection profiles for use by these initiatives.

### System Backup

There are two main strategies to follow when performing a system backup: one for the workstation level and the other for the network level.

### Workstation Strategy

The best backup strategy for workstations is to backup often. If the workstation is running the Windows OS, there are some simple backup tools already provided. There are also several utilities and programs available from reputable companies to aid users in performing backups. The following features can make backup chores more bearable: incremental backup, unattended scheduling and easy, simple restoration. Incremental backup saves changes made since the most recent full or incremental backup. This is important because users who do not

want to wait to backup a system can use incremental backup as a substitute for a lengthy full backup. Scheduling uses software automation to execute backup chores without the need for personal interaction. While the user must select and put in place a backup media, the user does not need to be present for the actual backup. Zip© drives and small tape drives are also cost effective solutions used to backup workstation data.

## Network Strategy

The best backup strategy for networks is an approach that combines several features to save time, effort and still assure complete backups. Execute full backups often. Since backups take up network, server and/or workstation resources, it is best to run full backups when nobody is working. Also, open files are skipped during backup and do not get backed up at all until some future time when the file is closed and not being used. Having few to no users holding files open will ensure the greatest backup saturation possible. Full backups are most efficiently executed in the evenings. Store the full backup tape off site. On each of the remaining workdays of the week, using a separate tape for each day, run an incremental backup and store it offsite, too. The last full backup of the month should be permanently moved off site and held for archival purposes. So, if a network is attacked by malicious code, these backup techniques will ensure data integrity and allow all systems to be recovered.

## Security Awareness Training

Security awareness is usually a first line of defense for an organization. Organizations should implement a security awareness-training program sanctioned by a recognized information systems security authority such as National Institute of Standards and Technology (NIST). An acceptable security program should be able to inform users about the threats of e-mail attachments, simple physical security, and protection of authentication mechanisms. The threats are much more numerous than these examples but statistical information indicates most users knowing very little about these threats.

## Configuration

Proper system administration is one of the best mechanisms to limit the number of vulnerabilities that can be exploited. CERT and other organizations publish vulnerabilities and fixes for those vulnerabilities. Every organization should be aware of the latest security patches and fixes for their equipment.

## Privacy Concerns

Organizations may own the intellectual property of employees and may also legally restrict computer activities to only those approved by management. A common practice is to present this warning to all computer users as part of the normal login message. This does not mean that ALL managers in an enterprise own ALL of the transactions of ALL of the employees. Especially unclear is how to handle the conflict that arises between privacy and monitoring. Use of IDSs and system-monitoring tools requires caution. Legal issues pose a potential problem to the deployment and use Detect and Respond technologies. As noted in NTIB#1, legal and regulatory issues are very complex and the “legal system has not yet made authoritative judgements on the issues.” The report illustrates the conflicting views on the

subject noting that “intrusion detection systems are sometime viewed as intrusive themselves, and . . . the position is taken that all information systems are subject to arbitrary monitoring at any time.”<sup>3</sup>

Sniffers that search for key words in messages (e.g., “attack”, weakness,” or “confidentiality”) as a standard set of *watchwords* may find key words used in an appropriate manner depending on the type of correspondence. Audit trail reports may contain full command strings (including parameters). The results of an analyst’s investigation of traffic patterns or traffic content within or interfacing to an enterprise (either in response to a possible intrusion or during an investigation following an attack) could be considered an unwarranted invasion of privacy. Activating and directing a potential adversary to a honey pot (deception server) raises privacy issues as well. It is important to refer privacy concerns to the appropriate legal and policy organizations for the enterprise prior to deployment and use of these technologies.

### 8.2.5.7 Technology Reference Model

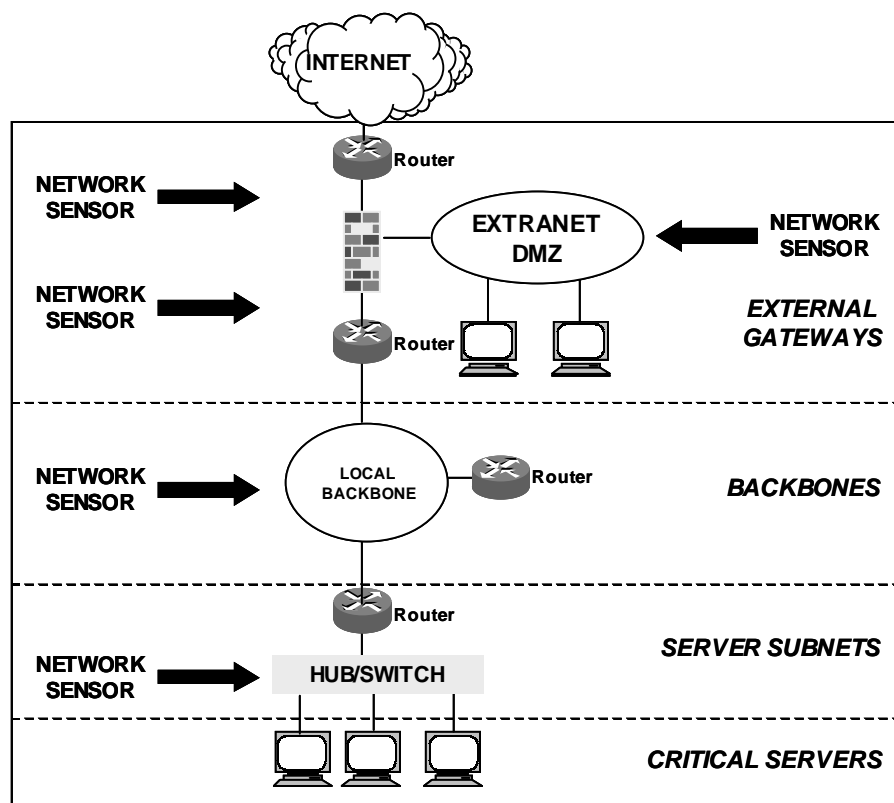
As discussed earlier in this section of the Framework, the Detect and Respond infrastructure is hierarchical by its nature. There is a tight coupling between the physical structures (of the local computing environment, enclave boundary, and system infrastructures), the processes that need to be performed, and the technologies that are available to realize those processes at each layer of the hierarchy. A technology reference model for this system infrastructure highlighting these relationships is provided in Figure 8.2-17.

The shaded areas of the figure represent a typical local environment (computing environment and enclave boundary). As discussed in earlier sections, the local environment is the natural location for host and network-based sensors (e.g., IDSs and vulnerability scanners). If detect and respond technologies (e.g., honey pots) are used, they are also located at this level of the hierarchy.

---

<sup>3</sup>National INFOSEC Technical Baseline – Intrusion Detection and Response.” Lawrence Livermore National Laboratory and Sandia National Laboratories, December, 1996., as reported in “Network Intrusion Detection and Response, a Technology Forecast, by William L. Cameron, AlliedSignal Technical Services Corporation, August 1998.





**Figure 8.2-17 Detect and Respond Technology Reference Model**

The processing above the sensors can be placed at every level of the hierarchy. Local environments have the option to deploy any and all aspects of processing and analysis, usually focused for their specific operations. Similar structures may also be available to focus at organizational, enterprise, and national levels. There is a decision-making capability needed at each level to interpret the operational implications of current situations and provide direction on courses of actions. This is typically performed with some collaboration at levels higher and lower as appropriate.

The network infrastructures that typically connect local environments together also provide the basic connectivity of these environments to various elements of the Detect and Respond infrastructure. This connectivity is needed to provide reporting up the hierarchy and information associated with response coordination back down.

The very nature of the reference model highlights the importance of selecting technologies that can interoperate with each other across the overall Detect and Respond infrastructure. While not shown, to realize a system infrastructure that can deal with an appreciable sized enterprise, that integration should extend into the system and network management infrastructures as well.

## 8.2.6 For More Information

The list of reference materials used in the preparation of this section provides an excellent base of knowledge from which to draw on relevant technologies, there are a number of additional sources of information. This section of the Framework focuses on on-line sources since they tend to offer up-to-date information. These include the following:

### IA Technology Framework Executive Summaries

An important segment of the Information Assurance Technical Framework (IATF) is a series of “Executive Summaries” that are intended to provide summary implementation guidance for specific case situations. These offer important perspectives on the application of specific technologies to realistic operational environments. While these are still being formulated, they will be posted on the IATF web site <http://www.iatf.net/> as they become available.

### Protection Profiles

The International Common Criteria and NIAP initiatives base product evaluations against Common Criteria Protection Profiles. NSA and NIST are working to develop a comprehensive set of protection profiles for use by these initiatives. An overview of these initiatives, copies of the Protection Profiles, and status of various products that have been evaluated are available at the NIST web site <http://niap.nist.gov/>.

### 8.2.6.1 Independent Third Part Reviewers of Relevant Vendor Technologies

ICSA Net Security Page [www.icsa.net](http://www.icsa.net)

Talisker's Intrusion Detection Systems [www.networkintrusion.co.uk/](http://www.networkintrusion.co.uk/)

Network Computing – The Technology Solution Center  
[www.nwc.com/1023/1023f12.html](http://www.nwc.com/1023/1023f12.html)

Paper on CMDS Enterprise 4.02 <http://www.Intrusion.com/Products/enterprise.shtml>  
(ODS Networks has changed its name to Intrusion.com)

PC Week On-Line [www.zdnet.com/pcweek/reviews/0810/10sec.html](http://www.zdnet.com/pcweek/reviews/0810/10sec.html)

### 8.2.6.2 Overview of Relevant Research Activities

Coast Homepage – Perdue University [www.cs.purdue.edu/coast](http://www.cs.purdue.edu/coast)

UC Davis <http://www.gidos.org/>

UC Davis [seclab.cs.ucdavis.edu](http://seclab.cs.ucdavis.edu)

## 8.2.6.3 Overview of Selected Network Monitor Vendor Technologies

Axent Technologies [www.axent.com](http://www.axent.com)

cai.net <http://www.cai.net/>

Cisco Connection Online [www.cisco.com](http://www.cisco.com)

CyberSafe Corporation [www.cybersafe.com](http://www.cybersafe.com)

Internet Security Systems [www.iss.net](http://www.iss.net)

Network ICE [www.networkice.com](http://www.networkice.com)

# References

1. Information Assurance Technical Framework (IATF) <http://www.iatf.net>.
2. National Institute of Standards and Technology <http://niap.nist.gov/>.

# Additional References

- a. Amoroso, Edward, Intrusion Detection. Intrusion.Net Books. 1999.
- b. Escamilla, Terry. Intrusion Detection, Network Security Beyond the Firewall. Wiley Computer publishing. 1998.
- c. Northcutt, Stephen. Network Intrusion Detection, An Analyst's Handbook. New Riders Publishing. 1999.
- d. Snapp, Steven R., et al. A System for Distributed intrusion Detection. IEEE CH2961-1/91/0000/0170. 1999
- e. Balasubramaniyan, J. S., et al. An Architecture for Intrusion Detection using Autonomous Agents. COAST Technical Report. 11 June 1998.
- f. AXENT Technologies, Inc. Intruder Alert 3.5 IDS Review Guide, May 2000.
- g. AXENT Technologies, Inc. Everything You Need to Know About Intrusion Detection, 1999.
- h. Schneider, Sondra, et al. Life After IDS. Information Security Magazine. Volume 2, Number 9. September 1999.
- i. Graham, Robert. New Security Trends for Open Networks. SC Magazine. October 1999.
- j. SC Magazine. Intrusion Detection. June 2000.
- k. Information Assurance Technology Analysis Center (IATAC). Tools Report on Intrusion Detection. Defense Technical Information Center. December 1999.
- l. Maes, V. How I Chose an IDS. Information Security Magazine. Volume 2, Number 9. September 1999.
- m. 13. Concurrent Technologies Corporation. Attack Sensing, Warning, and Response (ASW&R) Trade Study Report Intrusion Detection System. Report No. 0017-UU-TE-000621. April 14, 2000.
- n. Information Assurance Technology Analysis Center (IATAC). Tools Report on Vulnerability Analysis Information. Defense Technical Information Center. March 15, 2000.
- o. Ulsch, Macdonnell and Joseph Judge. Bitter-Suite Security. Information Security Magazine. Volume 2, Number 1. January 1999.
- p. Concurrent Technologies Corporation. Attack Sensing, Warning, and Response (ASW&R) Baseline Tool Assessment Task Anti-Virus Trade Study Report. Report No. 0017-UU-TE-000623. April 13, 2000.

**UNCLASSIFIED**

Detect and Respond as a Supporting Element  
IATF Release 3.0—September 2000

- q. Department of Defense (DoD) Chief Information Officer (CIO) Guidance and Policy Memorandum No. 6-8510, Guidance and Policy for Department of Defense Global Information Grid Information Assurance
- r. National Security Telecommunications and Information Systems Security Policy (NSTISSP) No. 11. National Policy Governing the Acquisition of Information Assurance (IA) and IA-Enabled Information Technology (IT) Products. January 2000.

## Chapter 9

# Information Assurance for the Tactical Environment

Communicating urgent, time-sensitive, or life-and-death information over wireless links in a military or quasi-military tactical environment presents unique information assurance challenges. This section addresses the specific security concerns associated with tactical information systems, and points out critical technology gaps in today's tactical communications environment. The section highlights key tactical-specific issues in an effort to generate credible Information Assurance (IA) criteria, resulting in significant and positive impact on IA technology developed by industry.

The first part of this section focuses on a description of the tactical environment and the types of threats specific to this environment. The latter part of this section covers several IA issues facing tactical users. Current and anticipated requirements for IA solutions are drawn from these issues. Finally, each section identifies current technologies in development or production that may satisfy key IA requirements, provides framework guidance on recommended technologies, and identifies substantive gaps in currently available security solutions. This insight will help guide United States (U.S.) industry in developing security technologies to satisfy the needs of tactical users. It also will assist government users in understanding the range of security solutions available, and the manner in which these solutions might be used.

While some of the key technologies discussed here may have been mentioned in previous sections of the IATF, the unique requirements of the tactical environment warrant a separate discussion for the benefit of equipment developers, integrators, and warfighters. This section focuses exclusively on those issues in which the tactical environment presents unique requirements for IA technologies. Tactical users should refer to other sections of this IATF for guidance on common IA technologies such as firewalls, virtual private networks, and intrusion detection systems.

This chapter of the IATF will be useful to the following types of organizations.

- U.S. Department of Defense (DoD) and commercial engineering support organizations responsible for design, integration, and life cycle support of tactical communications and information processing equipment.
- Military and other DoD organizations involved in conducting tactical operations.
- Other non-military organizations involved in tactical operations; e.g., law enforcement, Alcohol, Tobacco, and Firearms (ATF), Drug Enforcement Agency (DEA), the Coast Guard, emergency responders, search and rescue units, and Immigration and Naturalization Service (INS), other agencies involved with National Security/Emergency Preparedness (NS/EP) communications.

- Anyone whose operations are mobile or who has heavy reliance on urgent, time-sensitive, or life-and-death information often communicated over RF links.

## 9.1 Target Environment

### Definition of Tactical

In this context, “tactical” communications refers to a set of systems, products, and infrastructure that transfer time- and content-sensitive communications between wireless nodes, or from wired to radio transmission environments. These systems are typically used in military-style operations and require specific frequency allocation and spectrum management to avoid electromagnetic interference with commercial/civil communications.

The following set of characteristics are used to define tactical.

- Military-style operations.
- User-owned (or leased) equipment and infrastructure.
- Radio communications in licensed frequency bands.
- Communications in a hostile physical and RF environment.
- Classified or Sensitive But Unclassified (SBU) communications.
- Time-sensitive communications.

Because of the unique nature of the tactical environment, certain types of attacks are more common than others. Previous sections of this framework divide attacks into four categories: passive, active, insider, physical and distribution. Tactical forces place a high degree of trust in individual unit members and in the communications systems available for their mission. However, as the information transport network in tactical environments is often Radio Frequency (RF)-based, the potential still exists for an adversary to gain access to internal tactical communications systems, masquerading as an authorized user. The adversary then has the potential to conduct an “insider” attack. Thus, tactical communications systems must also defend against these types of attacks.

The majority of tactical communications systems are subject to both passive and network attacks by highly sophisticated adversaries, often with abundant resources. While not a tactical site attack, the recent Kosovo conflict demonstrated an increase in the sophistication of our adversaries. Individuals sympathetic to the Serbian forces attacked several U.S. and North Atlantic Treaty Organization (NATO) Web sites. Although most of these were denial of service attacks against publicly accessible web pages, more complex and malicious attacks can be anticipated in future conflicts.

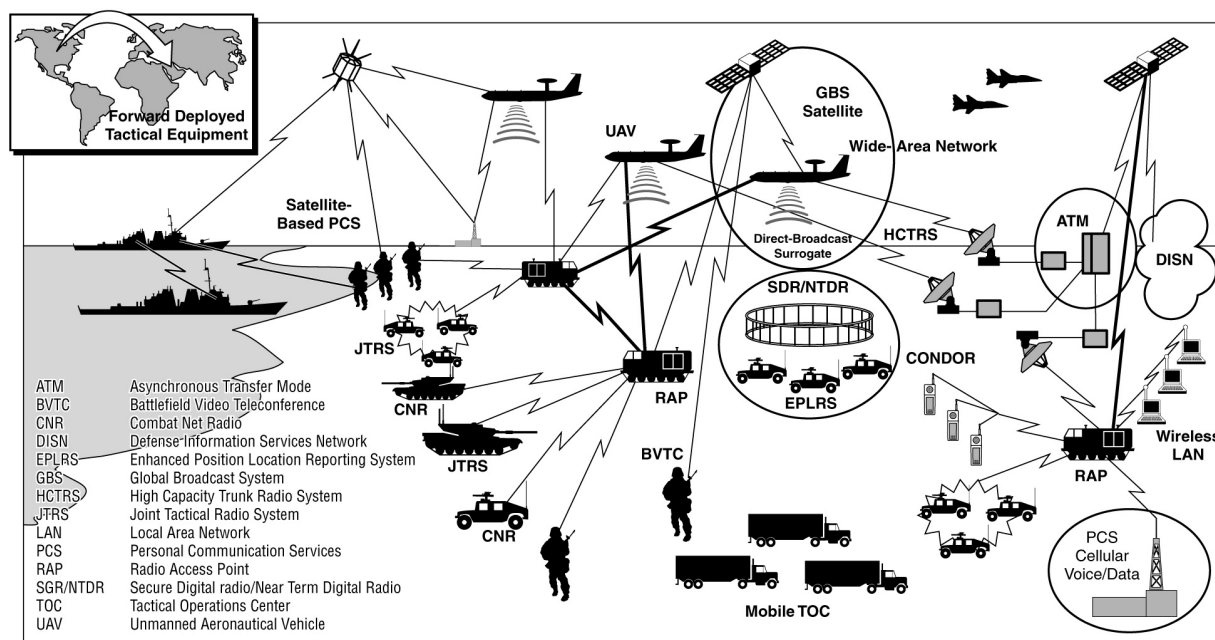
As noted in Section 5.2 (Wireless Networks Security Framework), commercial wireless users and service providers are often concerned with theft of service attacks. However, tactical users of military wireless communications systems are concerned with more destructive attacks threatening lives, or the national security, or both. Specific attacks that many tactical users want to prevent include:

- Geo-location (determining location of operators, confidentiality),

- Detection and interception of communications traffic (confidentiality),
- Jamming communications traffic (denial of service),
- Communications traffic analysis (garnering knowledge of activities from patterns of communications usage),
- Network intrusion and associated masquerading attacks (integrity, false message insertion, password sniffing),
- Theft of sensitive/classified information (confidentiality), and
- RF Fingerprinting (association of a particular medium with a specific user; i.e. unit identification based on radio characteristics).

## Military Examples

Examples of tactical communications scenarios vary based on the specific missions and Services involved. Figure 9-1 illustrates the complexity of deploying a total-force tactical communications suite to a battlefield. The diagram also shows the warfighter's reliance on key access points (satellite links and Unmanned Aerial Vehicle (UAV) airborne communication nodes) used to access the larger communications infrastructure. Communications architectures for non-military tactical operations can have similar characteristics.



**Figure 9-1 The Tactical Communications Environment**

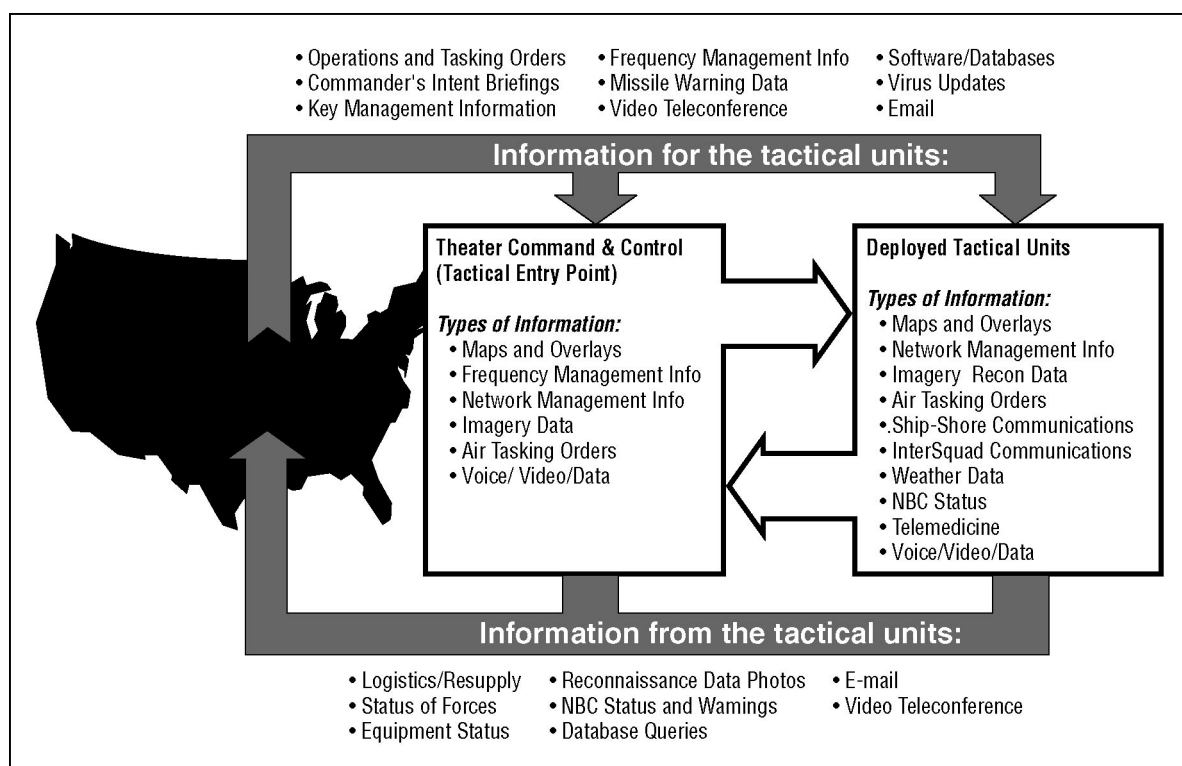
Clearly, not all of the systems shown in Figure 9-1 are interoperable, as the diagram may suggest. In fact, the majority of current tactical communications have a low degree of interoperability between Services. However, future systems like the Joint Tactical Radio System (JTRS) will provide increased interoperability between the Services' and allied networks, yielding an increased command and control capability for decision makers.



## UNCLASSIFIED

Information Assurance for the Tactical Environment  
IATF Release 3.0—September 2000

Figure 9-2 shows the types of information flowing into and out of a typical tactical environment to U.S. command sites. Major operational functions such as frequency management are often handled at a Main Operating Base (MOB) or command center, rather than on the front lines. Other functions provided from Continental U.S. (CONUS) locations include missile warning information from the North American Aerospace Defense (NORAD) Command, and Nuclear, Biological, and Chemical (NBC) fallout tracking from Los Alamos National Labs. These types of information pass back and forth between tactical forces and fixed locations in CONUS. Additionally, critical databases and imagery information are maintained either at the MOB or at the theater headquarters. Tactical units can access information on an as-needed basis, instead of bringing extra equipment to the front lines. Thus, a significantly large amount of data flows continuously between the main base in CONUS, the forward base, and forces on the front lines in a tactical scenario.



**Figure 9-2 Tactical Communications Information Flow**

Tactical communications are often defined by their environment and purpose rather than the specific equipment in use. In the past, tactical communications equipment was primarily composed of Government-Off-The-Shelf (GOTS) equipment. Such unique or “closed” systems, however, often require extensive support throughout their life cycles. In addition, it is often not cost effective to try to expand their capabilities to meet new requirements. Even with increased government budgets, the need for more capability has outstripped resources. Increased interoperability requirements and faster technological evolution have resulted in the increased use of commercially developed equipment in tactical communications. The trend in today’s tactical equipment design is to build open architectures where new advances can be added to systems efficiently.

A key example of DoD movement towards an open architecture is the JTRS. Recently, DoD identified the needs and benefits of combining various radio acquisition programs being proposed by the Services. As a result, DoD proposed development of a family of affordable, high-capacity tactical radios to provide line-of-sight and beyond-line-of-sight Command, Control, Communications, Computer, and Intelligence (C4I) capabilities to Warfighters. This family of radios will be capable of covering an operating spectrum from 2 to 2000 MHz, and will be capable of transmitting voice, video, and data. However, the JTRS is not a one-size-fits-all solution. Rather, it is a *family* of radios that are interoperable, affordable, and scalable. By building upon a common architecture, JTRS will improve interoperability by providing the ability to share waveform software and other design features between radios. The goal is to migrate today's legacy systems to systems compliant with the JTRS architecture. Section 9.8.3 (Technology Assessment) presents a more in-depth discussion of the JTRS.

The challenge of moving to an open architecture while remaining backwards compatible with existing legacy equipment and systems can seem overwhelming. Military systems have traditionally been designed for a specific type of environment, with little regard to future universal interoperability. However, tactical communications systems in the future will be required to inter operate effectively. For example, until recently, two separate devices were required if a commander wanted to place a call on a local cellular system and on a Satellite-Communications (SATCOM) link. Today, a single telephone will operate on both standard cellular and Low-Earth Orbit (LEO) satellite systems. Ideally, this same cell phone can then be integrated into other tactical communications networks like the Mobile Subscriber Equipment (MSE)/Tactical Packet Network (TPN) suite of equipment to maximize the operator's connectivity in the tactical environment, while minimizing the volume of equipment carried. To realize this vision, tactical systems will need to support common signalling plans and protocols such as Internet Protocol (IP) and Future Narrow Band Digital Terminal (FNBDT). Additionally, future systems such as the JTRS will handle multiple frequencies; multiple types of data (voice, data, and video); and multiple waveforms. Warfighters will drastically improve their situation awareness by accessing vital intelligence databases and imagery. Future tactical cellular systems and Personal Digital Assistants (PDA) will allow troops to pull down current satellite images or update enemy locations on the commander's map, giving the commander a better picture of the battlefield. These information advantages can only be realized, however, if the tactical information and communications systems possess sufficient levels of information assurance.

## Civilian Examples

Non-military organizations also employ systems that meet the tactical communications definition presented earlier. Examples include:

- First responders deploying to a terrorist incident.
- Communications support to the Secretary of State when she travels.
- Civil departments and agencies deploying to support missions under a variety of operational plans.
- Industry deploying network disaster recovery teams, cellular sites on wheels, and satellite telephone banks into disaster areas, as was the case in 1995 during the Hurricane Marilyn response on St. Thomas, VI.

## UNCLASSIFIED

Information Assurance for the Tactical Environment  
IATF Release 3.0—September 2000

A particularly interesting example is the new Florida Veterans Mobile Service Center consisting of a 43-foot mobile medical/dental clinic and veterans benefits. The center uses four cellular phone connections, two satellite links, and two laptop computers to link counselors with the state's Department of Veterans Affairs (VA) medical centers and benefits office, allowing them to access veterans' records and medical histories. Video-conferencing equipment allows VA physicians to interview patients directly from the mobile unit.

Probably the best example of non-military tactical operations is the Federal Emergency Management Agency (FEMA) in its role under the Federal Response Plan (FRP) as the coordinator of federal responses to Presidentially declared disasters and emergencies. FEMA coordinates FRP consequence management support to a number of national plans including the Federal Radiological Emergency Response Plan, the National Oil and Hazardous Substances Pollution Contingency Plan, and the FBI's Weapons of Mass Destruction Incident Contingency Plan.

As consequence manager, FEMA is responsible for organizing federal efforts to protect public health and safety, restore essential government services, and provide emergency relief to minimize the effects on the populace of a natural, technological, or terrorist event. To support the various operational facilities and teams that respond in accordance with the FRP, FEMA can deploy telecommunications assets from its six Mobile Emergency Response Support (MERS) detachments located in Massachusetts, Georgia, Texas, Colorado, and Washington, and its Mobile Air Transportable Telecommunications System (MATTS) located in Virginia.

MERS and MATTS assets can deploy to a disaster area to support federal, state, and local responders using a variety of communications transmission systems such as satellite, high-frequency, and microwave line-of-sight (LOS) interconnected by fiber optic cables to voice and data switches, local area networks, and desktop devices such as personal computers and telephones. Telecommunications can be provided for single or multiple locations within a disaster location. MERS and MATTS telecommunications assets can establish or reestablish communications connectivity with the public telecommunications system or government telecommunications networks and can interconnect facilities within the disaster region.

MERS and MATTS include these telecommunications transmission capabilities:

- Satellite. Ku-band satellite for quick connectivity that provides up to 48 lines for either telephones or data. International Maritime Satellite (INMARSAT) and American Mobile Satellite Corporation (AMSC) satellite terminals provide immediate single voice channel capabilities.
- Line of Sight Microwave. Microwave transmission to connect to the public network, provide connection to other facilities, or extend communications.
- High Frequency (HF) radio to communicate with federal, state, and local emergency centers via the FEMA National Radio Network and FEMA Regional Radio Network.
- Very High Frequency (VHF) and Ultra High Frequency (UHF) radio for local communications.

When deploying in to what amounts to a tactical situation, non-military organizations face some of the same IA issues and requirements as DoD. The requirements most important to non-military organizations are interoperability among response elements and protection from:

- Interception of communications traffic that is normally unclassified but may be sensitive.

- Denial of service.
- Network intrusion.

## Layout of the Tactical Communications Section

In order to adequately scope the key IA issues facing U.S. tactical forces today, representatives from the tactical community contributed to a list of the leading IA issues to be discussed in this section of the IATF. This list is certainly not all encompassing and may vary in order of importance for different users. However, the issues discussed here will apply to a variety of users and will highlight the IA deficiencies that exist in current systems. Joint and service-specific documents such as Joint Vision 2010 and the U.S. Army Warfighter Information Network document are used as key reference points for many of the tactical issues and requirements discussed in this section of the Framework. Unless otherwise noted, these issues are consistent with the issues described in the service's forward-looking documents.

The key IA issues identified by the tactical community that will be discussed in this section are:

- Wiping Classified Data From Tactical Equipment (9.2).
- Stored Data Protection In a Hostile Environment (9.3).
- Key Management In a Tactical Environment (9.4).
- Network Mobility/Dynamic Networks (9.5).
- Access to Individual Classified Accounts By Multiple Users (9.6).
- Secure Net Broadcast/Multicast (9.7).
- IA Solutions In Low Bandwidth Communications (9.8).
- Split-Base Operations (9.9).
- Multi-Level Security (MLS) (9.10).

Within each topic area, a brief overview is provided, followed by a discussion of IA requirements related to each topic. Tactical communications system users have critical equipment and infrastructure requirements over and above what the typical civil or commercial user requires. Anticipated requirements are added in the discussion to highlight requirement areas that will likely need to be addressed for tactical forces five to ten years in the future. These anticipated requirements are based on forward-looking documents such as Joint Vision 2010, the Concept for Future Joint Operations, and the Warfighter Information Network (WIN) Master Plan.

It is important to note that these anticipated requirements should not be considered “essential” for operations in a tactical scenario. Clearly, warfighters today employ technologies that do not meet many or all of these requirements. Rather, new technologies that incorporate these requirements would be better suited for tactical use than current systems. Thus, development of such technologies will improve the IA inherent in future tactical equipment and systems.

After the requirements discussion, relevant current technologies are addressed. Finally, each topic concludes with a section regarding Framework guidance. The guidance section presents technology recommendations for tactical users and Information System Security Engineers (ISSEs), and technology gaps highlight areas for future industry developments.

## 9.2 Wiping Classified Data from Tactical Equipment

### 9.2.1 Mission Need

U.S. military forces have been involved in an increasing number of non-traditional operations in recent years. Joint and multinational operations, peacekeeping missions, and support of FEMA efforts present challenges to the security of U.S. forces and systems that never before existed. During the same period, the U.S. military has adopted a host of new information and communications capabilities. Equipment formally used at the Secret or U.S. Only levels also is used for unclassified Federal Emergency Management Agency (FEMA) operations and in multinational operations. Furthermore, in recent years, nation states that were once on opposite sides of conflicts are now part of the NATO coalition forces. Thus, a new requirement has emerged to reuse tactical communications equipment at different classification levels for a variety of missions. IA technologies must be employed to provide a high degree of assurance that sensitive information used in one mission is completely wiped from the equipment before it is used in subsequent missions.

Tactical data wiping is typically done for one of three primary purposes: equipment storage, national level reuse, or multinational reuse. Residual classified or other sensitive information must be totally erased from any storage media residing in tactical communications or computer equipment. This includes information at several different classifications and handling caveats. The reuse of tactical communications equipment at different classifications applies virtually across the board to all types of equipment. In the past, systems such as the Secure Telephone Unit (STU)-III solved this problem by implementing a cCrypto-ignition Key (CIK) for each STU-III. The combination of a STU/CIK can be programmed to operate at any classification level. When the phone and the key are separated, they are each considered Unclassified/Controlled Communications Security (COMSEC) Item (CCI). Similar technologies are used in TACLANE and FASTLANE encryptors, as well as with the Krypton Personal Computer (PC) card in the tactical Secure Telephone Equipment (STE). However, creating these keys can take up to a week. Future use of programmable cryptography, multi-level security solutions (See Section 9.10.), and over-the-air updates for Type 1 cryptography will help alleviate this issue.<sup>1</sup>

For many years, tactical forces used communications equipment in what is referred to as a system-high environment. In other words, if the system handled information up to the Secret level, all equipment on the network was treated as Secret. Units often purchased multiple sys-

---

<sup>1</sup> Throughout this chapter (and other chapters and sections), reference is made to “Type 1” strength cryptography. In traditional usage, this has meant government-developed or sponsored equipment containing security mechanisms that meet some minimum strength of implementation. Enough assurance mechanisms were in place to reduce compromising failures to acceptable levels. In the context that the term is used here, Type 1 is generalized to include any source of equipment provided that robust minimums of cryptographic strength and assurance mechanisms have been included in the design. The exact definition of these assurances and strengths is beyond the scope of this document. This definition of Type 1 is also used in Section 5 (Defend the Network and Infrastructure).

tems to operate at different system-high classification levels. In some cases, declassification of equipment for reuse in another situation was possible, but generally time consuming. Declassifying equipment for use at lower classification levels will continue to take weeks, if not longer. When declassification is done before putting equipment into storage, the tactical user may be able to afford the extra time. However, if the equipment will be reused nationally or internationally, time may be a critical factor. In some cases, the declassification process may be overlooked entirely because of urgent mission requirements. With today's limited budgets, U.S. forces do not have the luxury of purchasing multiple sets of systems for each level of classification. Furthermore, the number of multinational operations that U.S. tactical forces are involved in has increased dramatically and will continue to increase in the coming years. Thus, finding solutions for this issue is vital. If IA solutions are not in place to enable rapid equipment reuse at different classification levels, tactical forces will be forced to purchase additional equipment for each system-high level or accept the risk that sensitive information will be compromised. The interim solutions of purchasing additional sets of equipment and relying on a time-consuming declassification process must be replaced by faster, higher assurance solutions.

As an example of multinational reuse of tactical equipment, recent NATO operations in the Balkans demonstrate the trend toward use of multinational forces in tactical operations. U.S. forces frequently report to coalition commanders from other nations. In addition to the usual issues (language, standard operating procedures) arising from a multinational chain of command, U.S. forces must protect cryptographic keys and algorithms from falling into the wrong hands, for a coalition partner today may be an adversary tomorrow. To prevent our IA solutions from being used against U.S. forces in the future, security solutions such as tamper-proof cryptography, programmable cryptographic chips, and over-the-air key load and zeroize functions should be implemented in future tactical communications equipment.

## 9.2.2 Consolidated Requirements

- IA technologies must be available to completely remove sensitive information from storage media on tactical communications and computer equipment and ensure that the data is not recoverable.
- These technologies must allow for equipment reuse at different classification levels.
- Equipment declassification processes must be accomplished rapidly (in a matter of minutes).
- Solutions such as tamper-proof cryptography, programmable cryptographic chips, and over-the-air key load and zeroize functions should be implemented in future tactical communications equipment.

## 9.2.3 Technology Assessment

To prevent our IA solutions from being used against U.S. forces in the future, security solutions such as tamper-proof cryptography, programmable cryptographic chips, and over-the-air key load and zeroize functions should continue to be implemented in future tactical communica-

tions equipment. A viable multi-level security solution, discussed in Section 9.10, (Multi-Level Security (MLS)) may also help address this issue.

For computer hard drives and other magnetic media, several software packages currently exist to purge classified data from a storage device. Two primary types of wiping software are available today: software that purges all data from a media, and software that purges deleted data from a media. These packages can also be used by certain tactical units to purge data from PCs and other magnetic media. However, much of the legacy communications equipment used by tactical units does not interface well with PC software or PC-based networks. Tactical radios may store sensitive information about a particular communications network that has to be erased before reusing the equipment in an unclassified scenario. Legacy cryptographic equipment can usually be zeroized with the press of a button, and new keys can be loaded at different classification levels. However, many of these legacy cryptographic systems are still considered sensitive even after they have been zeroized because of their internal design and the algorithms used. Newer programmable cryptographic chips will be able to wipe keys and algorithms from the chip, leaving a totally unclassified chip capable of being reloaded with new keys and algorithms.

With standard workstations, the weaknesses of current Operating Systems (OS) make the reuse of computers for different classification levels especially vexing. The allocation of data in swap files, the creation of temporary files, the storage of data in slack and unallocated space, and the actual non deletion of data despite using the delete command all constitute a potentially serious security hazard. Given the easy availability of hacking tools and forensic software, the possibility of data recovery is especially high. Although Commercial-Off-The-Shelf (COTS) memory “shredding” application software (BC Wipe, Erase, Kremlin, Puffer) exists—and there is a DoD standard for file wiping—the most secure solution is the total removal of all previously used storage media prior to reuse of the basic computer. This decision should be based on a careful risk analysis of the individual situation.

**Note:** Users should consult local security policy for a list of approved wiping software before using any of the software applications listed above.

## 9.2.4 Framework Guidance

Given the current state of technology, the best available solution continues to be removable storage media and zeroize functionality. Equipment can easily be reused in different missions by inserting a new storage media at the appropriate classification level. The zeroize function would also allow new cryptographic keys to be loaded at the appropriate classification level for the new mission. The desired solution involves the use of programmable cryptographic chips used in conjunction with a secure OS. The secure OS ensures all copies of sensitive files are handled at the appropriate classification level. Users without the appropriate authorizations cannot access the protected information. The programmable cryptographic chip would allow simple key and algorithm updates capable of upgrading or downgrading the equipment classification. Development and use of both programmable cryptography and secure OS are in their infancy. As technology matures, new solutions will be available to address this issue.

## 9.3 Stored Data Protection in a Hostile Environment

Tactical forces always have been faced with the possibility of enemy capture or overrun and the seizure of critical, sensitive, or classified information. In modern warfare, an increasing amount of information is stored electronically. While this has reduced the volume of sensitive documents and cryptographic material that must accompany a tactical unit to the battlefield, the problem of quickly destroying classified information in an overrun situation has merely changed—not been eliminated. Implementing strong, high-speed, and high-volume media encryption technologies would help mitigate the danger of compromised information, even if tactical communications or information system equipment falls into enemy hands. Alternatively, robust means of quickly rendering digital media unreadable are necessary.

The tactical requirement for media encryption differs from a non-tactical situation in two primary areas. First, the information stored in tactical equipment is often very perishable or time sensitive. That is, after a period of time, the utility of the information expires and it no longer requires protection. While this is not true for all tactical data, typically the media encryption needs to be only good enough to prevent the enemy from breaking the encryption within a short period (days to weeks). For example, information concerning an upcoming attack is classified only before the attack takes place. If information stored on a system pertains to an attack happening in three days, the encryption may only need to be strong enough to prevent an adversary from accessing the information for a week or more.

Second, tactical users often require extremely fast (near real time) media encryption. The media encryption process should be transparent to the tactical user, allowing the user to control the process in real-time and quickly protect the information in a time of crisis. If an Army unit is under attack by the enemy, a soldier may require the capability to rapidly encrypt large storage devices just in case the enemy captures the equipment.

### 9.3.1 Mission Need

Equipment subject to theft or recovery by an adversary must have the capability to adequately protect the information stored within the equipment. Current media and file encryption techniques are too slow for use in tactical situations. Media encryption of 1-2 GByte hard drives must be accomplished in a matter of minutes instead of hours. In tactical situations, zeroization is often used to destroy sensitive information if enemy forces will likely recover the equipment. Until strong, fast media encryption technologies are developed, zeroization will continue to be used in these situations. Once the equipment is zeroized, critical data is lost forever, and it cannot be recovered if the equipment is not captured. Thus, soldiers are often hesitant to hit the zeroize key if there is a chance of defeating the attackers. Unfortunately, this sometimes means that capture happens before zeroization.

Alternatively, sensitive information used in a tactical scenario could be maintained entirely in an encrypted state. Warfighters would then “pull,” i.e., decrypt, only the information needed at a particular time. The remainder of the disk or other storage device could remain encrypted until required by the warfighter, thereby limiting the amount of information that can be



recovered by an adversary. This method involves file encryption, instead of the more extensive media encryption technique that would encrypt the entire storage media. Thus, a method for pulling subsets of information from an encrypted drive while maintaining encryption for the remaining data on the drive is also a tactical requirement. This solution would enable encryption of the storage media that is transparent to the user because of the limited amount of information stored in the clear at any one point in time. Unlike zeroization, media encryption allows data recovery, enabling the soldiers to press the media encryption key first, so they can concentrate on defending themselves.

As stated previously, not all-tactical information is perishable. Some data stored on tactical equipment may require more extensive protection because the sensitive nature of the data persists beyond today's operation. Examples of these types of data would be information on weapons systems, classified procedures, or other information that will remain classified long after the tactical operation is complete. Clearly, the user must first determine the perishability of the information before deciding on the strength of encryption required to protect the data.

### 9.3.2 Consolidated Requirements

- Tactical communications systems subject to theft or overrun by an adversary must have a real-time method of protecting sensitive information. Tactical information is often time-sensitive or perishable. A decision must first be made about the perishability of the information. Then, the tactical user requires confidentiality services that can be rapidly applied to the information according to the sensitivity and perishability.
- A real-time means of protecting digital media must be available for the tactical user enabling the warfighter to quickly protect sensitive information in a time of crisis. Ideally, these services should operate transparent to the user.
- Near-term solutions utilizing file encryption must have a method for pulling subsets of information from an encrypted drive while maintaining confidentiality for the remaining data on the drive.

### 9.3.3 Technology Assessment

Tactical success requires encryption hardware and software that can meet time-critical requirements and provide real-time encryption/decryption. Tactical systems must process encryption requests at speeds essentially equal to those of unencrypted requests. High-performance, real-time bulk encryption requires data rates that stretch the performance parameters of available hardware and software. Media encryptors specifically protect the confidentiality and integrity of data storage media. They are designed to encrypt the entire contents of the storage media (less certain system files in computers).

Generally, tactical equipment that is subject to recovery and exploitation by the enemy is better protected by media encryption vice file encryption techniques. Much tactical information is time-sensitive and fast-moving. Sorting out information for file type encryption is not feasible; thus protection of the entire storage media is more desirable. This requires real-time media encryption to protect all the data in a timely fashion. The “wiring” of the battlefield down to the

individual soldier, and the enormous variety of communicated data, demands fast bulk media encryption and storage in a highly user-transparent manner.

Prime examples of the applications in the current technology are the developments in the FORTEZZA® family. Tactical applications for real-time encryption of mass storage devices including hard disks, floppy disks, tape drive, Compact Disc-Read Only Memory (CD ROM) and magneto-optical back-up storage are coming on line. Promising COTS developments in dedicated Protocol Control Information (PCI) card encryption accelerators and faster algorithms coupled with tamper-proofing technology need to be integrated in a total protection package to reduce the threat of exploitation of recovered/captured equipment.

### 9.3.4 Framework Guidance

In order to meet this requirement in the near term, rapid media encryption can be accomplished on a file by file basis, rather than a total media encryption basis. However, this method does not provide the desired degree of assurance that the OS has not made duplicate copies of sensitive information in temporary files. This Framework recommends further developments of trusted OSs, as well as faster media encryption technologies that will operate transparent to the user.

## 9.4 Key Management in a Tactical Environment

Overall key management for a tactical communication network involves generation, distribution, and storage of keying materials. Clearly, this process requires an extensive Key Management Infrastructure (KMI) to handle the number of users in a tactical environment. Fortunately, the U.S. military has spent many years improving the current KMI used to distribute symmetric keys to troops around the world. Entire documents have been written on the structure of the military's KMI. It is not the intent of this document to describe the entire key management process. Instead, this section will discuss some of the current issues related to key management in a tactical environment. These issues include black key transfer, remote rekey, transfer, zeroize functions, and key loading functions.

Remote rekey has become a major IA issue in recent years for several reasons. The capability of a user to rekey COMSEC equipment from a remote location eliminates the need to either bring equipment to a central location, or send key updates to field locations. Any dangers of key compromise along the shipping process are eliminated, along with drastically reducing the time required for key updates. More importantly in a tactical situation, if a node in a network should be compromised, a good network management and control system can lock out compromised nodes and remotely rekey all other nodes in a network. Thus, an adversary who obtains keys and communications equipment cannot listen to sensitive communications or attempt spoofing attacks against friendly forces by pretending to be a valid user on the net.

## 9.4.1 Mission Need

One of the primary concerns for the warfighter is the elimination of red key. The current Electronic Key Management System (EKMS) delivers black key from the Central Facility to the Local Management Device/Key Processor (LMD/KP). For the tactical Army, this brings keys down to the division level in a benign, secure fashion. However, transfer of keys from division down to brigade, battalion and below is performed by a soldier carrying a key fill device, such as the Data Transfer Device (DTD), full of red keys. This soldier is a target waiting to be exploited. Thus, the tactical warfighter requires a KMI with the ability to receive black keys all the way down to the end COMSEC unit. That is, there should be no point in the transfer of keys where they are stored red. This will minimize the risk of insider attack and ease compromise recovery.

Remote rekey and network management can be accomplished with Over-The-Air Rekey (OTAR) or across a landline, as with a STU-III or STE. Over-The-Air Zeroize (OTAZ) and Over-The-Air Transfer (OTAT) of keys are closely related to OTAR. These processes involve the rekey, zeroize, and transfer of keys across a communications link from a centralized key management center to deployed COMSEC equipment. One of the primary IA challenges with these processes is how to confirm the identity of the network control station and the end-user equipment. Without proper Identification and Authentication (I&A) services, a sophisticated adversary could conceivably impersonate the network control station, send out a key update, and essentially take control of part of the tactical network. Therefore, the first requirement is to implement high-assurance key management capability, utilizing remote rekey mechanisms in tactical networks to ensure access control, integrity, and confidentiality for the rekey message.

The second requirement for OTAR systems is an automated process for conducting OTAR that can run on any tactical automation system such as the Maneuver Control System (MCS). An operator at the key management center would program the software to automatically send out new keys at a designated time. Any system that doesn't acknowledge receipt is quickly identified by the OTAR system, and the status of that particular unit or individual would then be verified. These types of systems exist for the Digital Encryption System (DES) and other Type III federal systems but not for Type I tactical systems.

Third, a common key fill device is required to operate with multiple types of cryptographic keys and multiple end systems. If the tactical user requires three or four different key loading mechanisms in the field, units must bring extra COMSEC equipment to the field. With a single-key fill device this equipment burden could be drastically reduced.

Remote keying mechanisms are essential to eliminating the need to bring large numbers of COMSEC items to the field. For example, implementing OTAR and OTAT mechanisms, a unit would only need the initial key fill for COMSEC equipment deploying to the field. All other updates would be accomplished remotely. If tactical forces operating in hostile territory rely on remote keying, the chance of an enemy gaining access to COMSEC keys declines significantly. However, remote keying places a high degree of trust in the key management and network management functions. If tactical units rely on an automated system to send out key updates, significant IA must exist within the automated system. Tactical forces must have total confidence in the rekey process. Forward units must know that the enemy cannot spoof the network management station by sending out false COMSEC updates to friendly equipment. If there is

any doubt about the validity of keying information, units may choose to operate “in the clear,” without encryption, instead of possibly accepting a rekey from hostile forces.

Additionally, if any tactical COMSEC devices or keys are captured, all other nodes on communication nets using the compromised keys must be notified immediately. Many of these processes are in place today for single-key types in legacy cryptographic systems. However, the process is not as clear for Public Key Infrastructure (PKI) and reprogrammable cryptographic devices handling keys for multiple networks. Improvements in I&A of network control stations will provide a much higher degree of assurance that the enemy has not spoofed a network control station. A final requirement is the development of a KMI to deal with EKMS, PKI, and reprogrammable cryptography. Additionally, to fully realize the potential of programmable cryptography, current COMSEC algorithms should be integrated into programmable COMSEC chips.

## 9.4.2 Consolidated Requirements

- Tactical users require the development of a KMI to deal with EKMS, PKI and reprogrammable cryptography. High-assurance remote key management capabilities must be implemented in tactical networks, including methods for conducting OTAR, OTAT, and OTAZ. Additionally, processes must be established to disseminate compromised key information for PKI and reprogrammable cryptographic devices handling keys for multiple networks.
- Tactical users require a process to transfer black key all the way down to the end COMSEC unit on the battlefield, dramatically reducing the vulnerability of key compromise.
- High-assurance I&A services must exist for both network control stations and end users for OTAR, OTAT, and OTAZ.
- Tactical users must have an automated process for conducting OTAR that can run on any tactical automation system.
- Tactical users must have a common key fill device to operate with multiple types of cryptographic keys and multiple end systems.

## 9.4.3 Technology Assessment

This section will focus on technologies associated with key loading, remote rekey, OTAR, OTAZ, and OTAT. Also, a section on PKI has been added to address the movement to public keying in future DoD systems.

OTAR is not a new topic for tactical communications systems. The Army's mainstay radio system, Single Channel Ground and Airborne Radio System (SINCGARS), has a remote rekey capability. Other systems throughout DoD have this capability as well. The issues discussed in this section are specific to certain aspects of remote keying including: Type 1 automated tactical OTAR, Type 1 OTAZ, the development of a single-key fill device, and development of a common compromise policy and recovery method for programmable cryptography devices.

## UNCLASSIFIED

Information Assurance for the Tactical Environment  
IATF Release 3.0—September 2000

OTAR is an effective way to distribute key updates to deployed forces in a tactical scenario. It reduces the amount of keying material that must be transported to the field, which increases the risk of key compromise. Additional improvements to the OTAR process should focus on developing an automated process for conducting OTAR that can run on any tactical automation system (such as the MCS). An operator at the key management center would program the software to automatically send out new keys at a designated time. Any system that does not acknowledge receipt is quickly identified by the OTAR system, and the status of that unit or individual would then be verified. These types of systems exist for DES and other Type III federal systems, but not for Type I tactical systems.

In contrast to OTAR, there are very few approved OTAZ schemes for military radio systems. A common scheme should be developed for use in all future DoD tactical radio systems. Similarly, there is no single-key fill device currently available to support the variety of COMSEC systems fielded. With different key fill devices available for Type I, Type III, public key, and commercial key systems, a tactical unit often carries a multitude of fill devices to the field. A common fill device would lighten the load for the warfighter, and reduce the requirement to protect and store the additional devices. Some of the devices currently used for downloading keys to COMSEC devices include the DTD, the KYK-13, the KYX-15, or the KOI-18. The DTD is probably the most interoperable key loading device currently used, compatible with such COMSEC equipment as SINCGARS radios, VINSON, KG-84, and others that are keyed by Common Fill Devices (CFDs). The next version of the DTD, the DTD 2000, is currently under development.

Another requirement that must be met by a tactical key management system is a common compromise and recovery policy. If programmable cryptographic devices are used in tactical radios of the future, each unit may have radios keyed for multiple networks. The specific networks may vary from unit to unit or from one contingency to another. As an example, if a radio is compromised with keys for SINCGARS, HaveQuick, and Enhanced Position/Location Reporting System (EPLRS) nets, a chain of notification to include the designated key compromise authority for each type of key needs to be identified, including a set time for key changes and a new key distribution schedule.

## Public Key Infrastructure

Success in accomplishing the mission in the tactical environment depends to a large degree on the establishment of a secure means of moving information resources—data, voice, and imagery—to support the effort. Implementing a PKI will certainly not solve all tactical IA problems. However, a robust PKI could become a critical component of a fieldable IA solution for battlefield and other tactical operations.

PKI allows tactical users to interact with other users and applications, to obtain and verify identities and keys, and to provide other authentication services. There are three primary levels of assurance: high, medium, and basic. In the DoD, PKI certificates will be issued for medium and high assurance only. DoD has no plans to support a separate basic level infrastructure. This is not to imply that PKI services at the basic level of assurance will not be of importance to DoD, only that these services will be provided by the medium assurance infrastructure. High assurance is provided by Class 4 certificates such as FORTEZZA® cards. High assurance devices are generally hardware-based tokens providing protection for SBU mission-critical information over unencrypted networks (Type 2 information). Medium assurance refers to software

based end-user tokens (Class 3 certificates) requiring in-person or trusted agent registration that will eventually migrate to a common smart card such as the DoD identification card. Medium assurance certificates can protect less sensitive information such as support and administrative information. Basic assurance refers to lower assurance, software-based solutions providing minimal protection because of the lack of registration controls.

A critical issue for tactical communications is interoperability over a wide range of vendors' products and standards. This is compounded by the likely requirement to interoperate with a large number of PKIs from allied military forces and other elements of the U.S. and allied governments. These other PKIs may be based on different products, certificate policies, and algorithms. Technology in this area is still evolving. Key tactical issues such as compromise recovery, key recovery, and rapid personnel transfers must be addressed. Public key cryptography is one of the most promising emerging technologies, but the framework required to support a viable PKI needs to be carefully thought out and established.

## 9.4.4 Framework Guidance

Key management in a tactical environment has been handled by the Services for many years for symmetric key types. However, as the DoD moves closer and closer to adopting a total PKI solution, tactical key management will also require some modifications. This Framework strongly recommends that any new system under development have the ability to receive black key all the way down to the end COMSEC unit. In other words, there should be no point in the transfer of key where it is stored red. This will minimize the risk of insider attack, decrease the risk to the warfighter carrying red key, and ease compromise recovery. Current systems that provide an OTAR capability, e.g., SINCGARS, should continue to take advantage of their remote rekey functionality. As interoperability between networks increases, the Services must work to develop a common compromise and key recovery policy for use with tactical systems loaded with multiple COMSEC keys for different networks. This technology gap will be particularly important as tactical communications equipment begins to implement programmable Information Security (INFOSEC) devices. Furthermore, a single key fill device for all tactical COMSEC equipment does not currently exist. This is one area for industry to focus on in the near future. Finally, this Framework encourages the continued development of programmable cryptographic devices, and the implementation of current COMSEC algorithms on these devices. Future systems such as JTRS will play a key part in the use of programmable cryptographic devices, and in the refinement of current key management policies and procedures in the tactical arena.

## 9.5 Network Mobility/Dynamic Networks

U.S. tactical forces conduct a majority of their operations in locations outside the CONUS. Given this fact, there exists a need for these forces to maintain seamless network connectivity regardless of location. In the civilian world, a business traveler can remotely access his or her company's network from anywhere in the world through a dial-up remote access connection or by simply acquiring an Internet connection at the mobile location and accessing files and e-mail through a network connection. In either case, tracing phone numbers or IP addresses can trace

the travelling businessman to a specific location. Such location tracking is not desirable for a tactical user, because specific locations of tactical units are often sensitive, if not classified.

## 9.5.1 Mission Need

Consider the case of establishing a deployed Local Area Network (LAN) with an Internet server. A new host IP address must be assigned at each location, forcing frequent updates of the Domain Name Server (DNS). One requirement for mobile tactical users is the capability to seamlessly connect to a local subnetwork anywhere in the deployed tactical network. Tactical operations often combine equipment from different units, forming several different subnets. Users need continuous access to the network as they move between subnets, regardless of which unit “owns” the subnet. The tactical user does not have time to reconfigure local IP address information every time the subnet changes. Furthermore, IA technologies must exist to protect the packets against active and passive attacks by unauthorized individuals from both the home and foreign subnets visited by the tactical user. While making it easier for authorized users to travel between subnets, the deployed tactical network must still employ IA mechanisms that authenticate mobile users to prevent the adversary from gaining access somewhere in the network.

A different, but related, mobility requirement for tactical forces is the need for rapid setup and tear down of communications networks. Tactical network applications differ from fixed plant applications in that tactical networks are mobile. Tactical units rarely stay in the same location for the duration of an operation. Therefore, networks that require vast amounts of cabling are often impractical for use in a tactical operation. To the extent possible, bulky cabling should be replaced by wireless solutions in future highly mobile systems. Of course, wireless systems present additional challenges such as jamming and geo-locating that must also be addressed. The point is that security services should not increase equipment setup time for the warfighter. Secure wireless network solutions for tactical applications are a key area for industry development. IATF Section 5.2 (Wireless Communications) discusses wireless systems.

Tactical mobility can also be achieved by using global broadcast communications systems and UAVs used as communications nodes. Although these topics apply to tactical network mobility, they are covered more specifically in Section 9.7 (Secure Net Broadcast/Multicast).

A Tactical Operations Center (TOC) is today's central communications hub for most Army tactical information systems. Setting up a TOC and running all the required cabling can take 24-48 hours. This is too long. Therefore, rapid setup and tear down can become a major issue. An airborne unit may have more of a challenge with TOC mobility than a less mobile Army unit, since an airborne unit is a “shoot and move” unit, requiring a more mobile TOC. In this situation, full communications capability can lag behind the unit because of the time required to setup a TOC. Replacing cabling with wireless connections would drastically decrease setup time. Additionally, wireless solutions allow creation of a *mobile* TOC, installed in a set of three or four vehicles, with communications staying “up and running” while the TOC is on the move. The U.S. Army's First Digitized Division is attempting to implement a mobile TOC in several vehicles with wireless bridges and TACLANE encryptors. The TACLANE encryptor is discussed later in this section.

Regarding mobile networking, the security implications depend on the type of tactical application in question. Without dynamic networking solutions in place, seamless message addressing is more difficult. Individuals sending messages to tactical forces must know the network address of the recipient before sending a message. Also, an adversary may more easily locate U.S. forces at deployed locations by watching message headers flowing across a network. However, not all tactical units are particularly concerned about the enemy knowing their location. Thus, this issue will vary in importance depending on the particular tactical information system application.

Mobile wireless networks have an increased possibility of eavesdropping, spoofing, and denial of service attacks. The mobile networking concepts under development must account for information security hazards such as these in their development phase. For example, in an IP network, routers continuously broadcast routing tables to other nodes in the network to help other routers choose the best route to send IP packets. However, if this broadcast is done in the clear on a wireless net, an adversary could quickly glean an approximate picture of the layout of the tactical network. A second challenge in applying these technologies in the tactical arena involves incorporating routing and security functionality in smaller form factors such as handheld radios. Size, weight, and power requirements for computer equipment will continue to decrease as technology improves, which may help alleviate this issue. Future tactical equipment will require secure protection for over-the-air exposure of user information, addressing, system control information, and portable processing. Where routing functionality is provided in addition to the traditional radio applications, routing tables must be transmitted on a secure channel that all nodes in the network can access.

Finally, new mobile ad hoc networking technologies must remain backward compatible with certain legacy communications equipment. Even as new technologies become available, tactical units will retain much of their legacy communications equipment because of large upgrade costs and experience with current systems. Thus, legacy radio addressing will remain a key issue to consider when developing new mobile networking technologies.

## 9.5.2 Consolidated Requirements

- Tactical users must have the capability to maintain seamless network connectivity regardless of location or subnet. Network routing and domain name servers must have the ability to forward data to tactical users moving between networks. Users require continuous access to the subnets as they move through the field.
- IA protections for tactical networks must be flexible enough to operate on different types of equipment from various units around the world.
- IA solutions must prevent access to any subnet by unauthorized users.
- Many tactical users require protection against geo-location by an adversary. Therefore, dynamic networking solutions must provide confidentiality for specific location information where necessary.
- Tactical communications equipment must be capable of rapid setup and tear down, allowing greater mobility for the tactical unit. Security solutions should be applied in smaller form factors, e.g., handheld and man-portable.



- Mobile networking concepts developed for the tactical environment must address both passive and active attacks from a sophisticated adversary.
- Tactical wireless solutions should implement Low Probability of Intercept (LPI), Low Probability of Detection (LPD), and Anti-Jam (AJ) technologies to provide transmission security (TRANSEC) as required for the particular tactical mission.
- Advanced networking technologies must remain backward compatible with major legacy communications systems and equipment.

### 9.5.3 Technology Assessment

Significant advances in mobile IP technologies have made several of these tactical mobility requirements a reality. As discussed in IATF Section 4.4, Important Security Technologies, Internet Protocol Security (IPSec) used in mobile IP enables a mobile node to change its attachment point on the Internet while maintaining its IP address(es) as well as protecting their communications when visiting foreign subnets. Travelling between subnets resembles a cellular user roaming from one cell to another. However, future advances in mobile wireless communications will likely involve the use of the IP suite. Using IP in a cellular-like roaming situation creates several IA issues that must be solved.

The message originator wants assurance that a message will reach the correct destination, regardless of the physical location of the recipient, without any chance of interception or spoofing by an adversary. This must also be true even when the originator does not know the location of the recipient. Likewise, a recipient must ensure that received messages from the “commander” are indeed from the “commander,” regardless of where in the network the “commander” is located. In an attempt to solve these assured delivery and non-repudiation problems, a concept of “mobile ad hoc networking (MANET)” has been developed to support robust and efficient operation in mobile wireless networks by incorporating routing functionality into mobile nodes. Such networks are envisioned to have dynamic, random, multi-hop technologies that are likely composed of relatively bandwidth-constrained wireless links. This vision differs from Mobile IP technologies in that the goal of mobile ad hoc networking is to extend mobility into the realm of autonomous, mobile, and wireless domains, where a set of nodes, which may be combined routers and hosts, form the network routing infrastructure in an ad hoc fashion.

### Mobile IP and MANET

MANET is an autonomous system of mobile routers and associated hosts connected by wireless links. The routers are free to move randomly and organize themselves arbitrarily, thus allowing the network's wireless topology to change rapidly and unpredictably. Such a network may operate in a standalone fashion or may be connected to the larger Internet. [1] These nodes principally consist of a router, which may be physically attached to multiple IP hosts or IP addressable devices. This router may have potentially multiple wireless interfaces each using various wireless technologies. [1]

Mobile nodes are mobile platforms that make up a MANET. These nodes may be located on airplanes, ships, trucks, and cars. The MANET system may operate in isolation or may have gateways to interface with a fixed network. The MANET system consists of dynamic topology.

With this topology, nodes are free to move arbitrarily; thus the network topology, which is typically multi-hop, may change randomly and rapidly at unpredictable times and may consist of both bi-directional and unidirectional links. The decentralized nature of network control in MANETs provides additional robustness against the single points of failure of more centralized approaches. [2] MANETs also have limited physical security. Mobile wireless networks are generally more prone to physical security threats than cable networks. There is an increased possibility of eavesdropping, spoofing, and denial of service attacks with wireless networks.

This protocol permits mobile internetworking to be done on the network layer; however, it also introduces new vulnerabilities to the global Internet. First, the possibility exists for an adversary to spoof the identity of a mobile node and redirect the packets destined for the mobile node to other network locations. Second, potentially hostile nodes could launch passive/active attacks against one another when they use common network resources and services offered by a mobility supporting subnet. The first vulnerability can be surmounted by the strong authentication mechanisms built into both basic Mobile IP and route optimized Mobile IP. [2] By using PKI, a scalable countermeasure against the spoofing attack can readily be deployed. Currently there is an effort under way to surmount the second vulnerability.

Mobile IP and mobile nodes have several requirements to allow for maximization of security. First, when a mobile node is on its home network and a Correspondent Host (CH) sends packets to the mobile node, the mobile node must obtain these packets and answer them as a normal host. However, if the mobile node is away from its home network, it needs an agent to work on its behalf. [3] The second requirement is that of the expectation of the mobile nodes to retain their network services and protect their communications when they visit foreign subnets and the expectation of the foreign subnets to protect their network resources and local traffic while they are visited by the mobile nodes. A mobile node roaming over the Internet should have safe and persistent IP connectivity that is permitted by the policies of its home and visiting subnets. Persistency of IP connectivity means that the connections should be handed off quickly and correctly so that the mobile node can maintain its Transmission Control Protocol (TCP) sessions when it changes its network attachment point. [4]

More information on Mobile IP is available at Web site: [http://www.hpl.hp.com/personal/Jean\\_Tourrilhes/MobileIP/](http://www.hpl.hp.com/personal/Jean_Tourrilhes/MobileIP/). [2] For more information on MANET, visit Web site <http://www.ietf.org/html.charters/manet-charter.html>. [1]

## TACLANE/FASTLANE/TACLANE Internet Security Manager

In an effort to overcome some of the drawbacks and interoperability issues with current bulk encryption technologies, two Type 1 IP and Asynchronous Transfer Mode (ATM) encryptors have been developed for the National Security Agency: TACLANE (KG-175) and FASTLANE (KG-75). These encryptors provide access control, authentication, confidentiality, and data integrity for individuals or groups of users. TACLANE encryptors are more likely to be used in a tactical scenario because of size and mobility issues. The Army's First Digitized Division uses TACLANE encryptors with a wireless bridge to set up a wireless tactical operations center among a suite of vehicles.

The TACLANE encryptor will secure communications in a dynamic TPN, in the Defense Information Systems Network, or over the Internet, facilitating integration of these and other mobile and fixed networks. This encryptor operates at 45 Mbps for ATM networks and 4 Mbps for

IP networks. A new, smaller version of the encryptor, “TACLANE Lite”, is a PC card size device that is compatible with TACLANE. The PC card version supports data rates from 1 to 45 Mbps. The reduced size, weight, and power will allow greater operational interoperability.

These encryptors support different levels of secure transmission by employing crypto-ignition keys, much like a STU-III or a FORTEZZA card in the STE. When the CIK is removed, the encryptors are Unclassified/CCI. As mentioned in Section 9.2 (Wiping Classified Data From Tactical Equipment) changing the assigned classification level of a CIK is possible, but it requires a significant amount of time (potentially several days). Ideally, future systems will be able to operate at multiple security levels without going through a lengthy rekey process. The real strength of these encryptors comes from the integration of the TACLANE Internet Security Manager (TISM) in the tactical network. The TISM allows remote management of encryptors and their protected devices from a central location.

The TISM provides remote rekey of the FIREFLY keying material in the TACLANE and FASTLANE encryptors, reducing the chance of compromise by eliminating manual distribution of keys. Also, FIREFLY and traditional keys can be assigned to FASTLANE and TACLANE ATM virtual circuits with the ability to activate and deactivate them. Furthermore, audit data from encryptors throughout the network can be collected and reviewed in a central location, looking for errors or evidence of electronic attack on the network. A TISM operator can specify alternate TISM managers as a backup. If a TISM site is compromised or overrun, network management can be conducted from an alternate location. Future enhancements to the TISM include remote zeroization capability and electronic distribution of access control lists.

## 9.5.4 Framework Guidance

Until secure, wireless network solutions are implemented, tactical units will continue to use copper and fiber connections to connect local network nodes. Minimal security challenges arise using copper and fiber instead of wireless. The major drawbacks are longer equipment setup and tear down times and larger lift requirements because of the weight of the cabling. On the other hand, there can be a greater risk of jamming and geo-location when using wireless solutions. Thus, tactical wireless solutions should implement LPI, LPD, and AJ TRANSEC as required for the particular tactical mission. System integrators for tactical organizations should also pay close attention to the continuing developments in the mobile networking arena. Many lessons can be learned from the Army’s First Digital Division as they implement mobile wireless networking technologies and TACLANE encryption devices. Dynamic addressing schemes will also play a key role in improved communications for mobile users.

Additionally, Personal Communications Systems (PCS) on the battlefield are currently in the form of small lightweight cells. This allows the tactical user limited mobility in the Division and rear areas. PCS radio access points and cell sites need to be small and rugged enough to be mounted on vehicles that travel with the tactical users. These cells would have to operate with little or no operator involvement and the mobile networks would have to be self-configuring as the mobile cells move with respect to their users.

## 9.6 Access to Individual Classified Accounts by Multiple Users

Information systems often make use of shared directories or databases that can be accessed by a group of users for a specific purpose. Users expect to have individual e-mail accounts for sending and receiving messages, files, and other critical information. However, military and other tactical units tend to operate more as a group focused on a particular mission. When communicating with a unit, messages are sent to a particular position, or function within that unit, e.g., Commander or First Sergeant, as opposed to being sent to some specific individual by name (role-based access control versus individual access control). Unfortunately, this means a higher risk of messages or data ending up in inappropriate hands. This is a key concern if an insider threat exists within a unit. With recent advances in access control technologies, significant limitations can be placed on who (by name or by role) may access a particular account, file, or database. Thus, the danger of message traffic ending up in inappropriate hands is eliminated. These access control technologies work well in the commercial world, but it is not clear how well they transfer to tactical operational environments.

Communications systems of the past typically used role-based access control mechanisms, partially due to the lack of sophisticated individual access control technologies, and because of the need for accessibility by several operators on different shifts. Today, the standard password controls can be used in concert with other technologies such as biometrics (fingerprint, retinal, or iris scanners), PKI mechanisms (hardware and software), or other cryptographic tokens. Some of these methods present unique IA issues regarding access to information by a limited number of individuals. These potential solutions are discussed in more detail later in this section. The network must have the ability to uniquely recognize each individual in a tactical scenario and allow that individual access to information in accordance with their role-based need to know.

### 9.6.1 Mission Need

In a tactical scenario where a commander or other key individual could be replaced, captured, or killed, the chain of command is defined so the next person in the chain will assume command seamlessly. If Commander A is removed from the picture, Deputy Commander B must be able to assume command and have access to all messages and files that Commander A had. If the Commander is the only person with the “key,” and is captured, the Deputy Commander cannot effectively make command decisions because of a lack of information. Similar single points of failure may exist with system administrators or other critical positions.

As illustrated in the above scenario, new users/commanders must have the ability to access the same message and database capabilities as former users/commanders. Without the proper multi-user access control technologies in place, one of two outcomes will result. Either the unit will choose not to use the access control mechanisms for the communications equipment, or the unit will risk not having access to critical information should the key authorized individuals be unavailable. Therefore, tactical information systems require a fieldable network access control mechanism with the ability to uniquely recognize each individual in a tactical scenario and

allow that individual access to information and system use capabilities in accordance with that required and authorized for their role.

In the past, tactical units have typically chosen to use either widely disseminated passwords that are rarely changed, or no access control mechanisms at all. Physical security controls governed who had access to specific information or message services. Unfortunately, enemy capture of equipment has occurred, and enemy forces often became adept at using captured equipment to impersonate U.S. forces on U.S. radio channels. As a result, complicated and burdensome authentication schemes were devised to defeat these impersonation attempts. It is difficult to say how successful these authentication schemes were. Current tactical communications systems increasingly relay data without operator intervention and must rely on more sophisticated access control systems that provide a high degree of assurance regarding authentication of distant ends. Tactical users must make split second decisions that could have grave consequences. If the user suspects the distant end access control has been breached, messages received over the network will not be trusted. Furthermore, any new access control mechanism to be fielded in a tactical environment should be simple and reliable enough to assure the user that information is secure. As with any new technology, new tactical communications networks must earn the user's trust before reaching their full potential.

## 9.6.2 Consolidated Requirements

- Access control services on tactical equipment must be flexible enough to uniquely recognize each individual and allow that individual access to information is based on clearance level and current mission needs.
- Any access control mechanism must be simple and reliable enough to operate in a tactical environment, and assure the user that authentication information is secure.

## 9.6.3 Technology Assessment

IA solutions for this issue continue to develop rapidly. As stated in the Mission Needs section above, any solution must be able to uniquely identify users and grant them access to information in accordance with their individual clearance level. Possible solutions include implementing smart card technology on DoD identification cards, maintaining and using biometric information on all individuals involved in tactical situations, or assigning public key certificates to all DoD personnel reflecting authorized security levels. PKI solutions are described in Section 9.4.3 (Technology Assessment for Key Management in a Tactical Environment). Other technologies are described below.

### DoD-wide Certificates

DoD is currently planning to issue public key certificates to all military personnel for identification and encryption purposes. By direction of the Deputy Secretary of Defense, all DoD users will be issued, as a minimum, Class 3 (medium assurance) certificates by October 2001. Beginning in January 2002, the Class 3 certificates will be replaced by Class 4, high assurance certificates for all DoD users. [5] DoD PKI medium assurance certificates located on smart cards or floppy disks are starting to be used by DoD personnel interfacing with the Defense Fi-

nance and Accounting Service (DFAS). These certificates could transfer well to a tactical network application to validate the identity of system users and the authenticity of messages received from those users. The primary reason certificates exist is to associate a person with their public key. [6]

## Biometrics

Biometrics is the statistical analysis of biological observations and phenomena. Biometrics identity verification systems use biometrics as a method for recognizing a person by measuring one or more specific physiological or behavioral characteristics, with the goal of distinguishing that person from all others. Biometric devices must be based upon a characteristic that differs in a measurable way for each user. There are many characteristics that meet this criteria including iris scans, hand geometry, deoxyribonucleic acid (DNA), and fingerprints.

The application of biometric technology in fast moving tactical situations offers some clear advantages. Tokens, smart cards, and physical keys can be lost, stolen, or duplicated and passwords easily forgotten or observed. Only biometrics bases I&A on an intrinsic part of a human being—something that is always available and totally unique.

Applications are coming into use in the commercial and the civilian sectors of the federal and state government. Current military applications, to date, are sparse and appear to center more on use in fixed facilities as opposed to purely tactical applications. However, as the technology progresses, several tactical applications are likely to arise for biometrics. Much of this is anticipated because biometric devices are expected to become widespread in the commercial and government sectors in the next few years. Although biometric applications have been available for many years, recent reductions in the cost of biometrics devices and the introduction of new applications, i.e., controlling network login, Web server access, media encryptor access; are driving the deployment of biometric devices. Current shortfalls in the technology related to a tactical environment are:

- Lack of Standardization – Currently, the Government and commercial industry are working together to define a standard for biometric products. The Biometrics Application Program Interface (BAPI) will allow products from multiple vendors to interoperate, preventing one-vendor solutions. Products adhering to the BAPI standard are expected in the near future.
- Environmental Conditions – Environmental conditions in a tactical environment may reduce the effectiveness of some biometrics devices. For example, heavy rain may effect facial scanners, dirt or injuries may effect fingerprint scanners, or loud noises may effect vocal recognition devices. These conditions can effect the accuracy of the biometric devices. The use of biometric devices by tactical users wearing protective garments such as gas masks must also be addressed.
- Computing Power – Advances in computing power and in biometrics recognition techniques have reduced the computing power required by biometric devices making biometrics more attractive and affordable for strategic environments. However, the low power, low computing power tactical user may not be able to perform biometric verifications in a timely manner.

Despite these current limitations, biometrics offer some interesting future possibilities for tactical applications. As biometric devices become transportable, the possible applications for a

tactical environment become feasible. For example, military units frequently shared equipment, databases, and directories. Access to individual files and databases must be restricted to authorized users only. Biometrics could provide the unique discriminator necessary to restrict access to the authorized user. Users could carry their biometric signature on a smart card. When they require access to a system, they would insert their smart card, scan their biometric trait and gain access to the system. Each user carrying their biometric on a smart card could provide a strong authentication mechanism that is transportable across multiple units.

## 9.6.4 Framework Guidance

Until DoD realizes the full implementation of DoD PKI, tactical units should continue to use the role-based access control mechanisms in use today. In situations where one password is shared between multiple users, system administrators should assign unique usernames and passwords to each individual to decrease the chance of password compromise even though each individual has identical access privileges. Advances in biometric authentication products may or may not prove useful in the tactical arena. ISSEs and system integrators should pay close attention to new developments in this area to determine what applicability they may have to tactical communications systems.

## 9.7 Secure Net Broadcast/Multicast

DoD, military, and civil agencies conduct a large number of operations that involve the use of tactical broadcast equipment. These operations can range from U.S. military troops actively involved in war to law enforcement officials conducting a drug raid or seizure. The term “secure net broadcast” refers to a networked communications system where all transmissions from any node in the network can be received by every other node. For voice communications, this network resembles the Citizens Band (CB) radios used in the trucking industry. However, in a tactical environment, broadcast transmissions must maintain confidentiality and integrity during transmission to prevent interception by an adversary. Similarly, multicast transmissions are directed at a subset of nodes in a network. From the early entry phases and throughout the lifetime of tactical missions, voice and data information must be broadcast/multicast to multiple nodes securely and accurately. The tactical equipment used in these exercises must allow users to move rapidly with flexible and survivable voice and data communications.

### 9.7.1 Mission Need

Traditional Land Mobile Radio (LMR) systems may not have the range to handle broadcast communications over a large area; other broadcast/multicast solutions may be required. There are several technologies such as CONDOR, UAVs, Global Broadcast Service (GBS), and PCS that exist to help reduce these vulnerabilities. These technologies provide point-to-multi-point security solutions for wireless communication systems. They also secure data broadcast/multicast by providing a high-bandwidth communications networking infrastructure. In addition, several of these technologies use direct broadcast satellite technology to prevent data interception or jamming.

As mentioned previously, voice and data broadcast/multicast in a tactical environment are subject to many vulnerabilities. Whether it is a military troop in a hostile environment engaged in war or a civil agency performing a drug seizure, operational data must be kept secure and accurate while in transmission from one point to another. During data broadcast/multicast, the data could be intercepted, altered or jammed if not adequately protected. Any of these vulnerabilities could result in fatalities. For example, in a tactical environment, troops/law enforcement officials attempt to remain undetected while executing the mission or exercise to prevent geo-location, insertion of false messages, or communications jamming, thus giving an adversary the advantage. Any of these threats could lead to disaster for any mission or exercise.

## 9.7.2 Consolidated Requirements

Tactical communications equipment must allow operators to roam over a wide area and still be able to receive and send secure broadcast/multicast data over the local infrastructure. Secure network broadcast/multicast systems include the following security services requirements.

- Tactical users, while on the move, must be able to send and receive voice and data information in a secure and undetectable fashion. The minimum acceptable data rates for voice broadcast is 2.4 kbps.
- During the broadcast/multicast of voice and data information, this information must be protected from detection and identification, transmission jamming, geo-locating, RF signal attacks, Infrared (IR) signal attacks, and message insertion and modification.
- Tactical communication equipment must be capable of performing rapid, secure broadcast/multicast of high-volume military information such as maps, intelligence data, weather reports, and air tasking orders.
- Tactical communications equipment must have improved filtering to combat interference and jamming that will require advances in DSP (Digital Signal Processing).

## 9.7.3 Technology Assessment

Various security technologies have been developed to improve secure voice and data broadcast and multicast. These security technologies help to reduce the vulnerabilities identified in this Secure Net Broadcast/Multicast section.

### CONDOR

The CONDOR Program provides security in wireless telecommunications systems to meet the communication security requirements of DoD, military, and civil agencies. CONDOR provides point-to-multi-point security solutions for secure network broadcast/multicast service using the FNBDT signaling plan to connect various communications systems including IS-95 (Code Division Multiple Access (CDMA)), Advanced Mobile Phone Service (AMPS) CypherTac 2000 and the mobile satellite systems of Iridium, Globalstar, and ICO. This signaling plan is also interoperable with the tactical and office STEs. CONDOR phones could prove useful as a broadcast voice solution for tactical commanders on the battlefield. Essentially, commanders could



have a mobile conferencing capability from any location within the tactical cellular network. For more information on CONDOR and its technologies visit the following site: <http://condor.securephone.net>. [7]

## Unmanned Aerial Vehicle

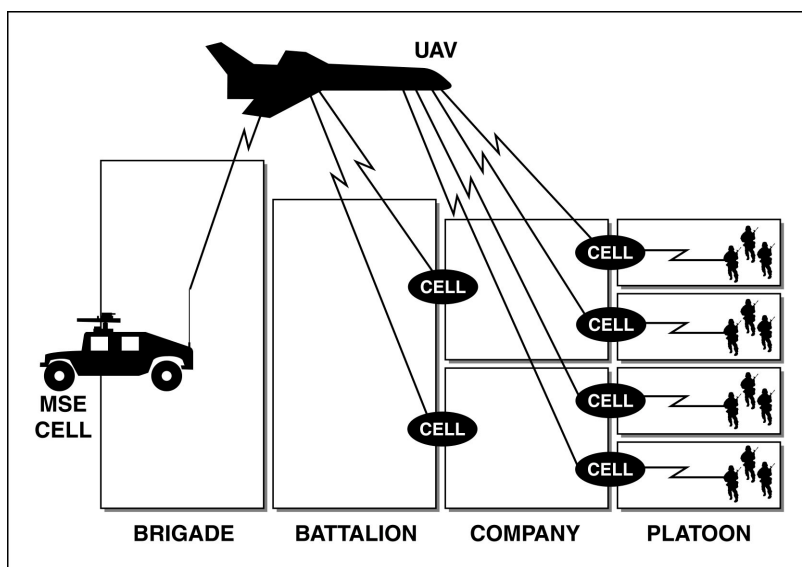
UAVs used as cell stations will help provide secure network broadcast/multicast communications for the tactical user. UAVs can provide a high-bandwidth, robust, and multimedia theater-level communications networking infrastructure that will protect net data broadcast/multicast from the vulnerabilities of jamming and interception. Currently, UAVs are used primarily as photo-reconnaissance platforms. However, to fully utilize the UAV on the battlefield, the UAV should be used as a cell station, or Airborne Communications Node (ACN). A tactical cellular network could be rapidly established by simply launching the UAV. From an altitude of 20,000 or 30,000 ft, an ACN produces a much larger cell area than a standard cellular tower. The UAV used as an ACN in the tactical Internet can provide warfighters with secure multimedia high-bandwidth Internet-type communications support in hostile tactical environments where communications must be broadcast and/or multicast to various destinations securely and accurately. For additional information on how UAVs can provide secure net data broadcast/multicast visit <http://www.darpa.mil>. [8]

## Global Broadcast Systems

GBS, developed by DoD, will increase the amount of national and theater-level information broadcast/multicast to deployed forces involved with operations in tactical environments. As the amount of broadcast/multicast data increases, GBS also provides increased security by using direct broadcast satellite technology. GBS enables commanders at the main operating base the capability to transfer vast quantities of information to forward units. This technology protects the data from vulnerabilities such as interception, jamming and modification. For more information on the security services of GBS, go to: <http://milsatcom.monmouth.army.mil/gbs.htm>. [9]

## Personal Communications Systems

PCS technology products have been developed to send and receive encrypted information from a portable PCS device to a tactical user of the Mobile Subscriber System. Tactical PCS secures network data broadcast and multicast by having radio access points or cell sites being made small and rugged enough to mount on vehicles that travel with the tactical users. For tactical missions that require data to be broadcast/multicast to users covering a large area, a UAV may be used to interconnect cell sites throughout the large area to keep the broadcast/multicast data secure. See Figure 9-3.



**Figure 9-3 Interconnecting Cell Sites using a UAV**

## 9.7.4 Framework Guidance

Future tactical systems will demand the use of commercial equipment and infrastructure. Thus, interoperable signaling plans and protocols should be integrated throughout all tactical systems. The FNBBDT is a network-independent, common cryptographic and signaling protocol that is implemented in CONDOR and the tactical STE. Inclusion of these protocols in such systems as the JTRS would dramatically improve interoperability, reducing the suite of duplicate systems a tactical user must carry.

Another technology gap involves the use of UAVs as an airborne communications node for tactical cellular. Current military UAVs, particularly the Global Hawk, Dark Star, and Predator systems are used exclusively for aerial reconnaissance. Significant improvements in tactical command and control would be possible by expanding the UAV mission to include its use as an ACN.

## 9.8 IA Solutions in Low Bandwidth Communications

One certainty of future tactical communications environments is that the warfighters on the battlefield at the lower levels of the command structure will continue to have smaller bandwidths and lower data rates available to them than the higher echelons. Also, the soldier on the ground or the pilot in the air—the warfighter at the point of the spear—has significantly less carrying capacity available for additional equipment than do fixed facility organizations. These constraints of bandwidth and lift are key drivers when implementing viable IA solutions at the tactical level.

The combination of limited funding for GOTS IA solutions and improvements in the strength of commercial solutions will lead to military systems of the future relying more on commercial IA tools to provide adequate security services. Unfortunately, IA technologies such as network monitoring systems occupy additional bandwidth that cannot be used for actual communications. To meet the objective of integrating IA solutions into the battlefield, these tools must operate with low bandwidth communications systems at the warfighter level without a noticeable degradation in the speed or accuracy of critical mission data traffic.

## 9.8.1 Mission Need

DoD would like to implement commercial IA tools in its tactical communications systems to decrease costs while increasing security and interoperability with the sustaining base. However, current tactical systems are not equipped to handle these commercial tools. As reported recently in *Federal Computer Week*: “Tactical battlefield networks under development by the Army and Marines to support operations on future digitized battlefields have vulnerabilities, according to Maj Gen Robert Nabors, commander of the Army's Communications-Electronics Command. 'Army tactical battlefield networks', Nabors said, 'do not have the bandwidth to handle commercial [IA] tools.'”[10] Furthermore, current planners estimate that the bandwidth available to the tactical soldier will likely remain low (10s of kbps). Given these constrained bandwidths, tactical users cannot afford IA solutions that impose additional bandwidth demands. Therefore, there is a requirement to adapt current IA technologies to lower bandwidth applications.

IA solutions that require significant bandwidth are not likely to be employed in the bandwidth-constrained environment of tactical operations, leaving tactical units with no alternative but to continue to operate with low—or no—assurance solutions. Network monitoring systems and intrusion detection systems employed on a tactical communications network can be monitored from the main operating base, or other rear echelon location. However, these systems send monitoring data from the end-user equipment back to the monitoring station. Thus, valuable bandwidth is occupied by monitoring traffic, decreasing the amount of bandwidth available to the warfighter or other operator for vital mission data. Without these IA solutions, a unit's network traffic could be subject to undetected interception and decryption by adversaries, ultimately leading to mission failure and loss of lives.

## 9.8.2 Consolidated Requirements

- Tactical networks require implementation of low profile IA monitoring tools that utilize minimal network bandwidth.
- In the long term, tactical networks must increase available bandwidth from tens of kbps to tens of Mbps to handle sophisticated, commercial IA tools.

## 9.8.3 Technology Assessment

Legacy military communications and information systems have traditionally been “closed” systems, meaning that equipment is designed specifically for use in one system. This is in

## UNCLASSIFIED

Information Assurance for the Tactical Environment  
IATF Release 3.0—September 2000

contrast to the current philosophy of migrating to an open systems architecture. In the past, low bandwidth communications used symmetric keying systems to provide confidentiality, and few network monitoring applications were available to ensure network security. Systems were not interoperable, and tactical forces learned to work around the constraints associated with closed systems. As communications and information systems move to an open systems environment, radios and networks from the fixed plant to the tactical domains must include a full suite of IA solutions in order to remain effective for military operations.

Remote network management plays a large part in maintaining the security of tactical networks. Utilizing advanced network monitoring applications, a technical controller can remotely monitor the security of several deployed networks from a central location. Tactical equipment typically has less bandwidth and processor capacity than fixed plant equipment. Therefore, it is more difficult to implement commercial IA tools in tactical communications networks and equipment. Current battlefield networks do not have the bandwidth to handle commercial tools like network monitoring and intrusion detection tools. However, programs are underway that may make it easier to integrate commercial IA tools into tactical systems. Two major programs that will benefit from this integration are the JTRS and the Marine Corps End-User Terminal (EUT).

**Note:** The Joint Tactical Radio program applies to several issues in this Framework. To avoid duplication of text throughout each issue, JTRS will be discussed exclusively in this section.

Joint Tactical Radio System (JTRS) will be the next generation radio for U.S. military forces in the 21st Century. In fact, in a memo to the Service Acquisition Executives in August 1998, the Office of the Assistant Secretary of Defense for Command, Control, Communications and Intelligence (OASD C3I) suspended all other “efforts to initiate any contracting activity to develop and acquire any radio system to include software-programmable radio technology.” The JTRS Joint Program Office (JPO) is charged with the responsibility of developing a family of JTRS products having common architecture and designed to serve different operational environments. As of this writing, the JTRS JPO was in the Phase 1 process of selecting the architecture to use for the production of the first JTRS prototypes (Phase 2). Therefore, specifics on JTRS will not be available until later revisions of this Framework.

JTRS will be a family of radios that provide simultaneous multi-band, multi-mode, and multiple communications using existing and advanced data waveform capabilities to ensure the timely dissemination of battle space C4I and global navigation information. The JTRS software-defined radio design represents a significant paradigm shift merging the commercial computer and networking industries with the wireless communications industry. While these technologies may prove beneficial in the commercial industry, implementing IA technologies into a Software Defined Radio (SDR) presents several new challenges. High-assurance software components must be developed and certified to perform in a manner acceptable for Type 1 security. A major benefit of JTRS is the scalability of the architecture. For a tactical unit, a handheld form factor should prove useful in satisfying the need for a low-bandwidth secure solution.

Overall, the benefits of JTRS significantly outweigh any technology issues that arise. Because the JTRS architecture is flexible and relies on many COTS products, a single Joint Tactical Radio can be scaled to meet the needs of any tactical unit. Airborne, vehicular, man-portable, and handheld versions will be available for use in the tactical arena, providing secure and non-

secure voice, video, and data communications using multiple narrow-band and wide-band waveforms. Operators will be able to load and/or reconfigure modes and capabilities of the radio while in the operational environment. Techniques such as OTAR, OTAZ, and other key management services are employed to overcome several of the IA issues discussed in this Tactical Framework. As this program develops, future versions of this Framework will address JTRS in more detail.

## US Marine Corps End User Terminal

The EUT is a technology currently in the testing phase by the U.S. Marine Corps. The EUT provides low bandwidth, networked communications at the squad level. However, the system currently lacks available security solutions. During recent Urban Warrior exercises, the Marines tested an EUT vest, composed of a mini-laptop computer running MS Windows, Netscape, and SRI's INCON Common Tactical Picture (CTP) software. These vests use differential Global Positioning System (GPS) for positioning and wireless Ethernet to communicate with one or more wireless access points. The mini-laptops have two PC card slots that are used by the wireless LAN PC card, and also for cellular phone PC card adapters. Additionally, high bandwidth Wide Area Network (WAN) connectivity is provided to the CTP via Very Small Aperture Terminal (VSAT) SATCOM and/or leased T1 lines. Thus, all the squads of Marines can access the CTP, including video feeds, intelligence images, and real time-updates. The CTP is also available to helicopters, boat units, light armored vehicles, and reconnaissance forces in the tactical area.

To date, all the Marine Corps Urban Warrior exercises have been unclassified, thus, minimal work has been done as far as cryptographic and IA solutions to secure the EUT and CTP software. Early testing has focused on integrating commercial networking technologies onto the tactical battlefield. Future solutions will likely employ some of the same high-assurance software products under development for the JTRS program. For more information on commercial wireless LAN technologies, refer to the IATF Section 5.2.3 (Wireless LAN).

### 9.8.4 Framework Guidance

Tactical users are encouraged to implement network monitoring, intrusion detection, and other IA tools in battlefield and other tactical environment networks. The adversaries of tomorrow will have the network savvy required to attack tactical networks. Detection and prevention of network intrusions will go a long way to insure the security of sensitive communications. Meanwhile, this Framework encourages the development of higher data rate (100s of Mbps) systems available at the lowest warfighter level with enough processing power to implement COTS security solutions in a handheld and man-portable form factor.

## 9.9 Split-Base Operations

The term “split base” refers to the situation where a unit deploys from its home base to a forward-operating base in or near the battlefield. As the U.S. decreases the permanent presence of its military forces on foreign soil, the number of such split-base operations will continue to increase. In forward operations, it is preferable to bring along as little infrastructure as possible.

The goal is to maximize forward capability. One approach is to leave infrastructure “at home” and rely on communications links to tie the warfighter at the front to the infrastructure at home. However, units must retain the capability to deploy to any site around the world, bringing an entire suite of equipment to the battlefield that can operate securely, without relying on specific IA tools available at that site. While the proximity to the battlefield may vary according to the service in question, e.g., Air Force versus Army units, the IA issues relating to split-base operations will generally remain the same. IA concerns for split-base operations actually incorporate several other issues already discussed in this tactical section. However, specific IA issues relating to split-base operations are discussed here because of the importance of secure communications during these types of operations.

To better support split-base operations, the services have programs in place to upgrade the communications infrastructure of military installations worldwide. DoD has embraced the idea of “network-centric warfare,” where tactical, logistics, and intelligence information becomes as much a weapon for the warfighter as firepower. Joint Vision 2010 puts networks at the center of military strategy for the next decade. Each service has separate programs in place to upgrade and standardize the client/server-based local, metropolitan, and wide area networks throughout the DoD. These programs are discussed below in the technology assessment area.

Infrastructure upgrades will drastically improve the support for deployed tactical forces, providing the capability to transport high-volume, real-time Command and Control (C2), and intelligence data to support contingency deployments and split-base operations during peacetime and war. As a rule of thumb, when a unit (or part of a unit) deploys to a forward area, there is an immediate demand for secure, high-capacity communications back to the main base. Today, most Air Force squadrons will deploy to an existing airbase near the theater of operations where communications capabilities are already in place. However, this is not always the case for tactical ground forces. When a tactical Army unit deploys to an area that does not have an existing communications capability, technologies must be available to enable rapid setup of secure voice, data, and video communications systems, linking the deployed unit to the home infrastructure. As the networking infrastructure of U.S. bases improves, tactical units must have the capability to connect securely back to their home networks. Tactical forces will likely rely heavily on SATCOM and other wide-band systems to provide these secure communications between home base and the TPN forward.

An example from the WIN Master Plan is used to illustrate the split-base operation concept. Today's equipment does not provide for multi-level security over a single channel. Current security policy for the TPN mandates that all hardware be accredited for Secret high operation. (The exception to this policy is the tunneling of SBU Standard Army Management Information System (STAMIS) users via in-line network encryption (currently the Network Encryption Systems (NES)) through the deployed TPN. For specific guidance on tunneling of lower classification data over a classified system-high network, refer to Section 5.3.7 in System High Interconnects and Virtual Private Networks (VPNs).

Today's typical configuration, shown in Figure 9-4 taken from the WIN, calls for the use of firewalls at gateway points between network types and High Assurance Guards (HAG) between the Secret Internet Protocol Router Network (SIPRNET) and Non-classified Internet Protocol Router Network (NIPRNET). Figure 9-5 shows the objective configuration implementing MLS with FORTEZZA® or other programmable cryptography at each node. Tactical forces that connect to the TPN need the ability to wirelessly pull information from SIPRNET,

# UNCLASSIFIED

Information Assurance for the Tactical Environment  
IATF Release 3.0—September 2000

NIPRNET, or the Joint Worldwide Intelligence Communications System (JWICS) databases from their deployed location. Also, improvements to the network infrastructure will improve command and control in split-base operations. Furthermore, security services such as confidentiality, data integrity, and access control mechanisms become increasingly important for a commander communicating with forward-deployed tactical forces. These services must continue to be a part of the TPN infrastructure.

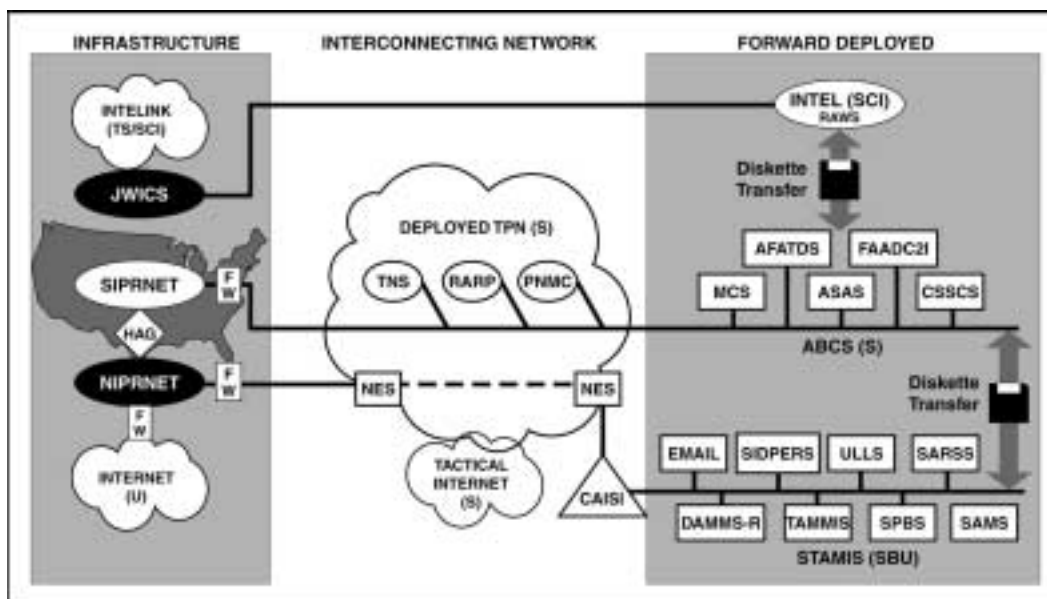


Figure 9-4 Near-Term Architecture [11]

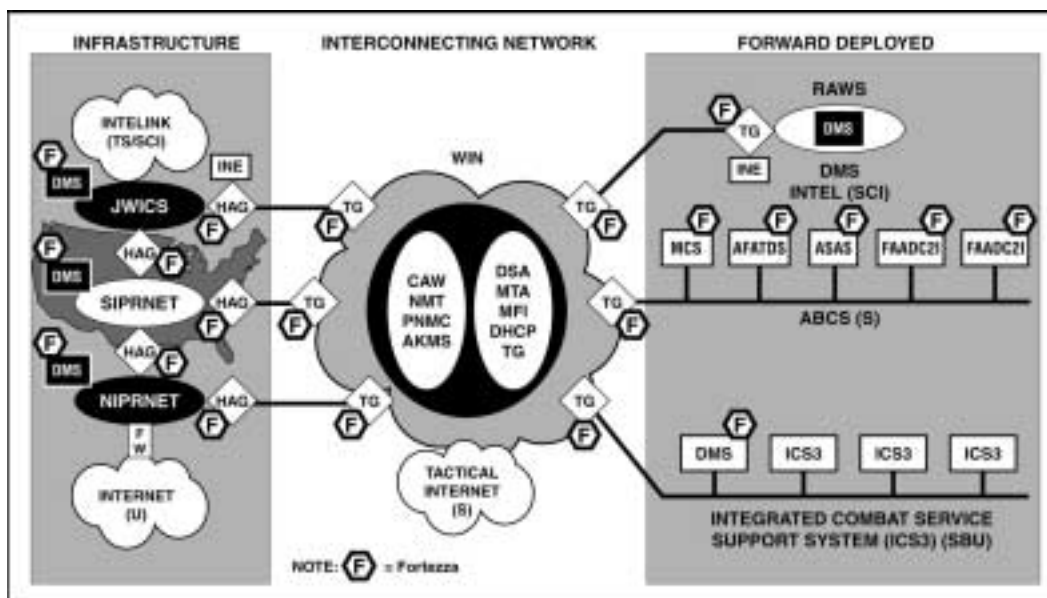


Figure 9-5 Objective WIN Security Architecture[11]

As stated previously, many of the IA issues discussed elsewhere in this chapter are particularly applicable to split-base operations.

## 9.9.1 Mission Need

Split-base operations are a culmination of all the tactical IA issues described in this Framework. Each IA issue must be addressed in order to securely execute split-base operations described in the WIN and other Joint Vision 2010 documents. Furthermore, as the number of permanent U.S. overseas installations decreases, the separation between “home” and “forward” will more and more often be between CONUS and “forward.” Network technology must provide a robust multimedia, theater-level communications networking infrastructure that can be rapidly deployed to support tactical operations. There are several security implications associated with maintaining communications links between the home base and a deployed location.

As an example, all types of information, from logistical supply data to intelligence data, traverses the communications link between the deployed location and the home base. For a sophisticated adversary with access to transcontinental communications, eavesdropping, disrupting, or denying the communications links necessary for successful split-base operations can give an adversary a significant military advantage.

## 9.9.2 Consolidated Requirements

The goal of a successful split-base operation is to maximize forward capability, while minimizing the amount of infrastructure required at the forward location. Thus, in addition to the requirements listed in the previous sections, the following requirements exist for IA in a tactical split-base operation:

- Infrastructure upgrades must occur in home-base networks to improve the support for deployed tactical forces. These upgrades must provide the capability to transport high-volume, real-time C2, and intelligence data such as battlefield video teleconferencing and transfer of satellite imagery to forward units.
- Tactical units must bring a suite of equipment to the battlefield that can be securely configured at any site, without relying on IA solutions available at that site.
- Technologies must be available to the warfighter at forward locations to enable rapid setup of secure voice, data, and video communications systems.
- IA technologies must be in place to prevent a sophisticated adversary from eavesdropping, disrupting, or denying the communications links necessary for successful split base operations. Proper implementation of security solutions discussed in Chapters 5 through 8 of this IATF can provide adequate protection for a split-base operation.

## 9.9.3 Technology Assessment

Well coordinated split-base operations require a sophisticated communications infrastructure at the base level in the CONUS. Based on guidance from Joint Vision 2010, the services have



## UNCLASSIFIED

Information Assurance for the Tactical Environment  
IATF Release 3.0—September 2000

kicked off several programs aimed at improving this infrastructure at the base level. These programs are discussed below.

The Navy has the Information Technology for the 21st Century (IT-21), which defines a standard, networked computing environment, based on commercial technology, for its ashore and afloat units. Key Army initiatives include the Outside Cable Rehabilitation (OSCAR) program, the Common User Installation Transport Network (CUITN), the Army's DISN Router Program (ADRP), and the Digital Switched Systems Modernization Program (DSSMP). These programs will update the Information Technology (IT) infrastructure at Army facilities in the US, providing an all-fiber ATM network to support real-time wide-band data requirements like video teleconferencing. Finally, the Air Force is implementing a base-level Combat Information Transport System (CITS) that includes installation of fiber-optic cable, ATM switches, hubs, and routers at 108 bases. As a vital part of CITS, information protection hardware and software will be installed as part of an Air Force standard network management system.

### Theater Deployable Communications Integrated Communications Access Package Program: Rapid Communications Setup in a Drop-in Airbase

The U.S. Air Force has also contracted for the development of a new advanced rapid deployment communications network used to deploy critical communications assets at a "drop-in" airbase. The program, called the Theater Deployable Communications Integrated Communications Access Package Program (TDC-ICAP), will provide secure and non-secure voice, data traffic for local area, intra-theater, and inter-theater communications using commercial components. The deployment of the TDC-ICAP will enable all of the U.S. Air Force elements (command and control, intelligence, logistics, and mission support functions) to function in a coordinated manner from initial deployment through sustainment.

The TDC provides a ground to ground communications infrastructure designed to transmit and receive voice, data, and video communications securely to or from wireless, satellite, or hard-wired sources. The system is both modular and mobile, allowing the Air Force to tailor the system to its specific needs and to transport the system anywhere in the world. Thus, TDC-ICAP drastically reduces the communications problems typically associated with airlift and manpower. The system is configured into common man-transportable transit cases to optimize airlift capability and to ease the problem of ground deployment.

TDC-ICAP interfaces with legacy TRI-TAC equipment through an adaptation of existing SMART-T technology developed for the Milstar system. Additionally, the ICAP is compatible with the telephone systems in 39 countries around the world, providing connectivity through a commercial Private Branch eXchange (PBX) to the local phone system. The center of the TDC-ICAP complex is the Base Hub, which supports all users located at its specific location. Additionally, all off-base communication passes through the Base Hub for distribution and is handled by the Off-Base Hub for specific interfaces, bulk encryption and decryption, and multiplexing.

The TDC-ICAP provides secure, tactical communications services to forward-deployed Air Force units virtually anywhere in the world. Rapid deployment of a core communications capability is central to the success of this program. Core communications can be set up in 1.5

hours after the initial pallets of equipment are delivered on site. Access is provided for TRI-TAC KY-68 encryptors. Two-wire and Integrated Services Digital Network (ISDN) interfaces are available at all nodes in the system allowing connection of STU-III or STE terminals for secure voice and secure fax capabilities. In addition, it is designed for transition to Defense Message System (DMS) compatibility, when that system is phased in. [12]

## 9.9.4 Framework Guidance

Split base operations will continue to present new technological challenges to the tactical unit commander. As the communications infrastructure improves, the forward commander will have access to increased bandwidth and unparalleled connectivity to rear-echelon networks. Tactical units will have the ability to access both the NIPRNET and the SIPRNET from their forward locations. One of the key technology gaps identified in this framework involves pulling information from the SIPRNET over a wireless link. A commercial PDA user can pull a map off the Internet, get directions, or access a database at the office from virtually anywhere in the country. However, a soldier on the battlefield has no way to access the SIPRNET to pull down a classified map or view overhead imagery. Continued developments in JTRS may help resolve this issue.

## 9.10 Multi-Level Security

As the U.S. military and other agencies with tactical missions move towards the next generation of radios and communications equipment, MLS has become an increasingly important technology hurdle. MLS implies a communications device that can simultaneously process data communications at different levels of classification. A radio on an unclassified network, e.g., HaveQuick in an Air Force network, will need to communicate with both unclassified networks, and data systems in a tactical internet operating at the Secret High level. Interoperability—the exchange of data between different classification levels—has become a necessity. As a result, MLS solutions are needed to integrate the majority of individual military communications systems into an interoperable ensemble of capability. Because of the difficulty involved with fielding a true MLS solution, this section focuses on MLS more as an objective than a requirement.

Traditional security policies mandate strict physical separation of systems and data at different classification levels. However, as the military moves towards a Software Designed Radio (SDR), physical separation is difficult, if not impossible, to achieve. MLS solutions will integrate high-assurance hardware and high assurance software solutions, eliminating the need for separate COMSEC devices and red processors at each independent classification level. Integrated MLS solutions yield critical size, weight, and power reductions, lightening the load for a tactical warfighter.

A cornerstone of multi-level security solutions is programmable cryptography. Programmable cryptography is a set of hardware and software capable of changing COMSEC algorithms and keys, allowing one device to interoperate with several different COMSEC devices. Current legacy communications equipment typically uses a COMSEC device particular to that equipment or to the specific channel on which a radio is operating using one COMSEC algorithm at a time.

In contrast, programmable cryptography enables communications equipment to load several different COMSEC keys simultaneously, allowing a single radio to “talk” on several different nets without requiring separate COMSEC devices or having to reload COMSEC for each net. Additionally, new algorithms can be added via secure software, and old ones can be deleted. Last, upgrades to programmable cryptographic devices are done in software, instead of hardware board replacements of legacy COMSEC equipment. This issue corresponds to Section 9.2 (Wiping Classified Data from Tactical Equipment).

### 9.10.1 Mission Need

True multi-level security solutions (at Type 1 security levels) have never been achieved for tactical systems. Communications at different security levels remains a complicated challenge. Separate red processors are required at each classification level, as well as separate buses and red devices for each level. Unfortunately for the tactical warfighter, this means more equipment in the field. A transition must be made from Secret-high operations to Multiple Independent Security Levels (MILS), and eventually to true multi-level security through the use of programmable cryptography.

A true MLS solution, as proposed in JTRS, would implement a programmable cryptographic chip in a single radio. Several different levels of cryptographic key would be loaded in the same chip, allowing the airborne troops to carry only a single radio into battle, freeing up part of their limited load for other items, such as ammunition. Use of programmable cryptography for MLS will increase interoperability between networks at different levels and decrease critical equipment requirements for the warfighter.

### 9.10.2 Consolidated Requirements

- Multi-level security solutions are needed to integrate the majority of individual military communications systems—increasing interoperability and reducing critical size, weight, and power requirements for the tactical user.
- A transition must be made from Secret-high operations to MILS, and eventually to true multi-level security through the use of programmable cryptography.
- Programmable cryptographic solutions used in concert with trusted OS must be available in the future enabling tactical communications systems to enable multiple levels of classified information on a single radio.

### 9.10.3 Technology Assessment

Multi-level security solutions will eventually be implemented in hardware or software or a combination of both. A hardware approach relies on physical separation of data at different classification levels, and it can be difficult to upgrade should modifications become necessary. However, by using a hardware-software combination solution, the hardware effects can be minimized. Hardware elements such as programmable cryptography can be used to eliminate the need for separate COMSEC devices and Red processors at each classification level. Part of

this section briefly discusses some of the programmable cryptography programs under development.

Additionally, a hardware-software combination MLS design may include use of a trusted OS, coupled with a trusted middleware solution. A high-assurance, software-based data control scheme ensures data separation for different classification levels. The advantages of this type of implementation are flexibility, portability, and minimal hardware dependency. Also, new security technologies can easily be added through software upgrades. A large number of real-time OSs are currently available. The choice of which OS to use for a particular application should be made judiciously, considering such issues as interoperability and performance parameters. Systems such as JTRS require Portable OS Interface Unix (POSIX) (IEEE 1003) compliance for the OS.

There are several major programmable cryptography programs currently underway, including AIM, Cornfield, FORTEZZA® Plus, Cypris, and the Navy's Programmable Embedded INFOSEC Program (PEIP). Certain devices fit better in different form factors, and allow several channels to operate simultaneously. Specific solutions should be chosen judiciously on a case-by-case basis. It is not the intent of this section to cover each program in detail or to recommend a specific device. Rather, in an effort to increase equipment interoperability and to decrease the amount of COMSEC equipment required in the field, this framework encourages continued improvements to current programmable cryptographic devices.

Programmable cryptography on embedded cryptographic chips will help pave the way to achieving full multi-level security solutions. Refer to the earlier discussion about JTRS for an example of a future tactical application of MLS. Programmable cryptography relies on high assurance components that essentially perform the function of maintaining separation of data at different classification levels. Instead of physical separation, these devices maintain strict data separation within the chip. Essentially, successful implementation of these chips in tactical communications equipment will reduce the amount of equipment required in the field and will reduce the number of COMSEC keys and equipment to be maintained in a hostile environment. Coupled with proper media encryption and zeroizing technologies, a true multi-level security solution will significantly enhance the effectiveness of tactical communications.

## 9.10.4 Framework Guidance

True multi-level security solutions do not currently exist. This Framework encourages continued research in the area of programmable cryptography, as well as in the development of trusted OS, to approach true MLS implementation. The JTRS program has a requirement for MLS operation three to six years down the road. Until that time, systems will operate with MILS. As a stepping stone towards MLS, MILS implies multiple classification levels of data in the same system as separate channels. Until true MLS is achieved, tactical units should implement MILS systems and components wherever possible to lighten the equipment load on the warfighter.

## 9.11 Additional Technologies

Given the format of this chapter, certain tactical systems that will play key roles in future tactical communications did not seem to fit any of the specific categories discussed above. Therefore, these systems are discussed here: Tactical STE, ISYSCON, and Battlefield Video Teleconferencing (BVTC).

### Tactical Secure Telephone Equipment

STE is the next generation of secure voice and data equipment for advanced digital communications networks. The STE consists of a host terminal and a removable security core. The host terminal provides the application hardware and software. The security core is a FORTEZZA® Plus Krypton cryptographic card, which provides all the encryption and other security services. The STE is available in two models: Office STE and the Tactical STE.

The Tactical STE provides secure tactical and strategic digital multimedia communications, interoperating with legacy TRI-TAC equipment, while also providing basic ISDN and STU-III compatibility in a single unit. The Tactical Secure Telephone Equipment (TAC/STE) provides direct connection to tactical communication systems in the field and offers full office features and connectivity for use in garrison. The design is based on the open, modular architecture, allowing efficient software upgrades to deployed units. The Tactical STE is Tri-TAC/MSE Interoperable and supports 16/32 kbps CVSD clear secure operation via LPC/CELP. In addition the Tactical STE PCMCIA Cryptography uses a removable FORTEZZA® Plus Krypton Card that supports SBU through Top Secret/Sensitive Compartmented Information (TS/SCI) traffic. For more TAC/STE information, visit <http://ste.securephone.net/>. [13]

### ISYSCON

Any tactical force deployment will require a number of communications networks. MSE, the mainstay for tactical area communications, operates alongside TRI-TAC assemblages. A vital flow of information gathered by the Joint Tactical Information Distribution System (JTIDS) is simultaneously relayed throughout the battlefield for air defense. Enclaves of soldiers will respond to urgent information passed over their Combat Net Radios (CNR). The EPLRS constantly updates and transmits its location information. The complexity and magnitude of these communications networks demand a means of integrating systems control to maximize the effectiveness and availability of the various systems and to ensure their interoperability. A program called ISYSCON provides this tactical area communications management capability.

The ISYSCON program brings a higher level of integrated communications management to theater tactical communications through a common mechanism, complete with automated tools, to seamlessly integrate communications systems at all levels. ISYSCON optimizes the application of standard Army Frequency Management, COMSEC, and Communications-Electronics Operating Instruction (CEOI) modules, provides automatic interfaces to the Battlefield Functional Area Control System (BFACS), and incorporates unique decision aides and embedded training capabilities.

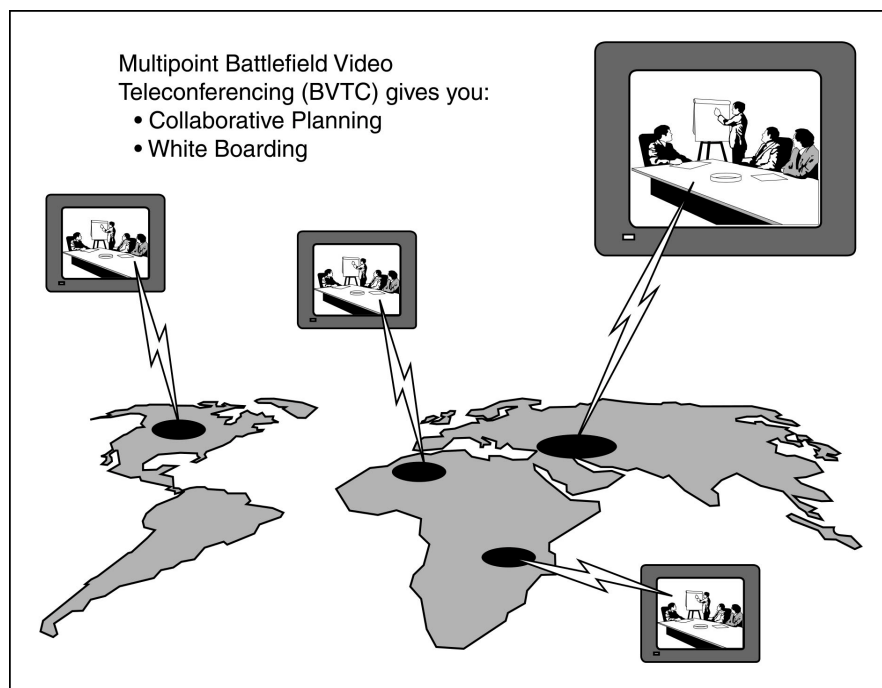
In the near future, joint communications planning and management for regional Commander in Chief (CINC's) and joint forces commanders will be provided by the emerging Joint Network Management System (JNMS). JNMS will facilitate communications network engineering, monitoring, control and reconfiguration. It also will perform frequency spectrum management and information assurance management.

## Battlefield Video Teleconferencing

Faster processor speeds and improved modulation techniques have boosted commercial use of VTC dramatically in recent years. Naturally, the desire to make use of this capability has transferred to the tactical battlefield. Battlefield Video Teleconferencing (BVTC) is a state-of-the-art, near full-motion interactive VTC system that enhances coordination and provides an additional combat multiplier to the warfighter. This technology can be applied at many levels through the battlefield. Two areas that will likely see great enhancements by the use of BVTC are warfighter C2 and telemedicine.

BVTC enhances C2 by allowing the warfighter to effectively disseminate orders, clearly stating his intent. The warfighter can conduct collaborative planning and whiteboarding with subordinate commanders and key staff elements. (See Figure 9-6.)

Medical units are supported by telemedicine from remote deployment areas, where skeletal medical forces receive assistance from specialists at sustaining-base hospitals. Other applications exist at several regional medical centers (Tripler, Walter Reed, and Landstuhl) to provide specialized diagnosis and care to remote medical facilities. Telemedicine will project the valuable expertise and skills of rear-based specialists to forward-deployed medics.



**Figure 9-6 Battlefield Video Teleconference**

Commercial development of VTC should drive the development of faster, highly capable network VTC applications. With the addition of data integrity and confidentiality mechanisms, VTC should transfer well to tactical applications. The wide bandwidth signals used with BVTC will require high-speed cryptographic solutions. A high-tech adversary could gain battle damage assessment or other sensitive information simply by intercepting a telemedicine BVTC channel. The development of high speed, reprogrammable cryptography will speed the implementation of the necessary INFOSEC solutions to BVTC.

**UNCLASSIFIED**

Information Assurance for the Tactical Environment  
IATF Release 3.0—September 2000

BVTC components (cameras, monitors, computers, microphones, etc.) are user-owned and operated. The features and capabilities employed at each echelon or activity will be based on the requirements of that specific echelon or activity. The Army's WIN architecture will provide the bandwidth and throughput required to support BVTC for both point-to-point and multipoint conferencing. BVTC capability will be provided to users of the WIN with nominal impact on the remainder of the network.

## References

1. Mobile Ad-hoc Networks (MANET) Web Page  
<<http://www.ietf.org/html.charters/manet-charter.html>>.
2. Mobile Ad Hoc Networking Working Group Web Page <<http://info.internet.isi.edu/in-notes/rfc/files/rfc2501.txt>>.
3. Hewlett Packard Web Site  
<[http://www.hpl.hp.com/personal/Jean\\_Tourrilhes/MobileIP/](http://www.hpl.hp.com/personal/Jean_Tourrilhes/MobileIP/)>.
4. BBN Technologies Web Site, <<http://www.net-tech.bbn.com>>.
5. Deputy Secretary of Defense Memorandum, Department of Defense (DoD) Public Key Infrastructure (PKI), May 6, 1999.
6. DFAS PKI Study March 10, 1999  
<<http://www.gradkell.com/PKI/DfasPkiStudy/DfasPkiStudy.PDF>>.
7. Condor Wireless Security Web Site, August 9, 2000 <<http://condor.securephone.net>>.
8. Defense Advanced Research Projects Agency (DARPA) Web Site, August 01, 2000  
<<http://www.darpa.mil>>.
9. Global Broadcast Service (GBS) Web Site, August 10, 2000  
<<http://milsatcom.monmouth.army.mil/gbs.htm>>.
10. Brewin, Robert and Daniel Verton. "DoD Leaders Mull Internet Disconnect." Federal Computer Week, April 19, 1999.
11. Warfighter Information Network Master Plan, Version 3. 3 June 1997.
12. Motorola Web Site <<http://www.mot.com/GSS/SSTG/ISD/ic/TDC.html>>.
13. Secure Terminal Equipment (STE) Web Site, August 10, 2000  
<<http://ste.securephone.net/>>.

## Additional References

- a. Network Security Framework, Version 1.1. December 3, 1998.
- b. Shalikashvili, Gen John M. Joint Vision 2010. Joint Chief of Staff: Washington DC, Jul.1996.
- c. United States Army Communications-Electronics Command. CECOM Vision 2010. New Jersey, 1997.
- d. JBC Information Assurance (IA) Tools for the Joint Task Force (JTF) Phase III Assessment Quicklook Report Summary. February 1999.
- e. Linton, Dennie. Global Broadcast Service: Shrinking the Year-2000 Battlefield by spreading the Word (Globally).
- f. Fillgrove, Ted. "Update on Enhanced Position-Location Reporting System"



## UNCLASSIFIED

Information Assurance for the Tactical Environment  
IATF Release 3.0—September 2000

- g. Newton, Harry. Newton's Telecom Dictionary. Telecom Books, October, 1998.
- h. Marine Corps Operation Urban Warrior Web Page  
<<http://www.defenselink.mil/specials/urbanwarrior/>>.
- i. SRI International InCON Webpage  
<[http://www.systech.sri.com/InCON\\_Description/coverpage.htm](http://www.systech.sri.com/InCON_Description/coverpage.htm)>.

# Chapter 10

## A View of Aggregated Solutions

---

This chapter will be provided in a later release of the Framework.

This page intentionally left blank.

# Executive Summary

## Chapter 1 – Introduction

The Information Assurance Technical Framework (IATF) document was developed to help a broad audience of users both define and understand their technical needs as well as select approaches to meet those needs. The intended audience includes system security engineers, customers, scientists, researchers, product and service vendors, standards bodies, and consortia. The objectives of the IATF include raising the awareness of IA technologies, presenting the IA needs of Information System (IS) users, providing guidance for solving IA issues, and highlighting gaps between current IA capabilities and needs. Chapter 1 outlines the information infrastructure, the information infrastructure boundaries, the information assurance framework areas, and general classes of threats. Then the Defense-in-Depth Strategy is introduced and the overall organization of the IATF document is presented.

## Chapter 2 – Defense-in-Depth Overview

When developing an effective IA posture, all three components of the Defense-In-Depth Strategy—people, technology, and operations—need to be addressed. This framework document focuses primarily on the technology aspects of Defense-in-Depth. The technology objectives and approaches in four focus areas, explained in the sections that follow, address the needs of the private, public, civil, and military sectors of our society.

Chapter 2 provides an overview of the Defense-in-Depth technology objectives and gives two examples of federal computing environments. The Defense-in-Depth objectives are organized around the four Defense-in-Depth technology focus areas:

- Defend the Network and Infrastructure
  - Availability of Backbone Networks
  - Wireless Networks Security Framework
  - System High Interconnections and Virtual Private Networks
- Defend the Enclave Boundary
  - Protection for Network Access
  - Remote Access
  - Multi-Level Security
- Defend the Computing Environment
  - End User Environment
  - Security for System Applications
- Supporting Infrastructures
  - Key Management Infrastructure/Public Key Infrastructure (KMI/PKI)
  - Detect and Respond

## Chapter 3 – Information Systems Security Engineering Process

Chapter 3 describes the systems engineering and Information Systems Security Engineering (ISSE) processes. The ISSE process is presented as a natural extension of the systems engineering process. Both processes share common elements: discovering needs, defining system functionality, designing system elements, producing and installing the system, and assessing the effectiveness of the system. Other systems processes: systems acquisition, risk management, certification and accreditation, and life-cycle support processes – are explained in relation to the ISSE process. Chapter 3 also provides suggestions on how the Common Criteria might be used to support the ISSE process. The processes described in this chapter provide the basis for the background information, technology assessments, and guidance contained in the remainder of the IATF document.

## Chapter 4 – Technical Security Countermeasures

This chapter of the IATF provides the background for detailed technical discussions contained in later sections of the IATF. It presents a general discussion of the principles for determining appropriate technical security countermeasures. The chapter includes a detailed description of threats, including attacker motivations, information security services, and appropriate security technologies. Using the methodology described in Chapter 3 (Information Systems Security Engineering Process), threats to the information infrastructure result in the identification of vulnerabilities followed by a managed approach to mitigating risks. Chapter 4 explains how primary security mechanisms, the robustness strategy, interoperability, and Key Management Infrastructure/Public Key Infrastructure should be considered in the selection of security countermeasures, technology, and mechanisms. These decisions form the basis for developing appropriate technical countermeasures for the identified threats, based on the value of the information.

## Chapter 5 – Defend the Network and Infrastructure

Chapter 5 describes the Defend the Network and Infrastructure technology focus area of the Defense-in-Depth Strategy. The chapter describes the types of network traffic—user, control, and management—and the basic requirements to ensure that network services remain both available and secure. Organizations that operate networks should defend their networks and the infrastructures that support their networks by establishing clear Service Level Agreements (SLA) with their commercial carriers that specify metrics for reliability, priority, and access control. Organizations must recognize that their data may be unprotected during transmission and take additional steps. Chapter 5 describes current strategies for defending networks (including data, voice, and wireless) and the corresponding network infrastructures.

## Chapter 6 – Defend the Enclave Boundary/ External Connections

Defense of the enclave boundary in Chapter 6 focuses on effective control and monitoring of the data flows into and out of the enclave. Effective control measures include firewalls, guards, Virtual Private Networks (VPNs), and Identification and Authentication (I&A)/access control for remote users. Effective monitoring mechanisms include network-based Intrusion Detection System (IDS), vulnerability scanners, and virus detectors located on the Local Area Network (LAN). These mechanisms work alone, as well as in concert with each other to provide defenses for those systems within the enclave. Although the primary focus of boundary protection is on protecting the inside from the outside, protected enclave boundaries also use technology and mechanisms to protect against malicious insiders who use the enclave to launch attacks or who facilitate outsiders gaining access through open doors or covert channels. The technologies discussed in Chapter 6 include firewalls, guards, virus/malicious code detection systems, IDS, and multi-level security systems. The IA strategy for defending an enclave boundary should flexibly implement those policies governing communications both between secure enclaves and between secure enclaves and external systems. The IA strategy must also provide the management capabilities for verifying compliance with policies governing defense of the enclave boundary.

## Chapter 7 – Defend the Computing Environment

Chapter 7 discusses the third technology focus area of the Defense-in-Depth Strategy, Defend the Computing Environment. The computing environment includes the end user workstation—both desktop and laptop—including peripheral devices. Servers include application, network, web, file, and internal communication servers. A fundamental tenet of the Defense-in-Depth Strategy is prevention of cyber attacks from penetrating networks and compromising the confidentiality, integrity, and availability of the computing environment information. For those attacks that do succeed, key are early detection and effective response to mitigate the effects of attacks. Intrusion detection, network scanning, and host scanning are the measurement functions that, on a continuous or periodic basis, determine the effectiveness of the deployed protection systems. Chapter 7 also addresses host-based sensors including those that operate in near real time as well as those that operate off-line.

## Chapter 8 – Supporting Infrastructures

Supporting Infrastructures is the fourth technology focus area of the Defense-in-Depth Strategy. The IATF addresses two supporting infrastructure entities: KMI/PKI and Detect and Respond. KMI/PKI focuses on the technologies, services, and processes used to manage public key certificates and symmetric cryptography. The discussion concludes with recommendations for the features needed to achieve the three Global Information Grid defined assurance levels: basic, medium, and high. The Detect and Respond section of Chapter 8 addresses providing warnings, detecting and characterizing suspected cyber attacks, coordinating effective responses, and performing investigative analyses of attacks.

## Chapter 9 – Information Assurance for the Tactical Environment

The tactical environment, in which military or military-style operations are conducted, presents unique information assurance challenges. In this operational environment, there is heavy reliance on the communication of urgent, time-sensitive, or life-and-death information often over wireless links. . In the past, tactical communications equipment was primarily composed of Government-Off-The-Shelf (GOTS) equipment. Decreased budgets and increased interoperability requirements in today's military organizations have led to the increased use of commercially developed equipment in tactical communications. Included in this use of commercial equipment is the use of commercial wireless networks and equipment in the tactical environment. Chapter 9 discusses the information assurance needs of the tactical environment, highlighting key tactical issues and identifying the associated security implications.

## Chapter 10 – A View of Aggregated Solutions

This section of the Framework is included in recognition of the fact that the needs of most users are represented not by any single technology focus area, but by some combinations of them. A future release of the Framework will include a discussion of developing and evaluating security approaches that are aggregations of the recommendations from the individual categories.

## In Closing ...

This Framework document is principally intended as a reference document to provide insight and guidance to security managers and system security engineers into how to address the information assurance concerns of their organizations. It is tutorial (vice prescriptive) in nature in recognition of the fact that many organizations face unique challenges that don't lend themselves to "one size fits all" solutions. This document offers insights intended to help improve the community awareness of the tradeoffs among available solutions (at a technology, not product level) and of the desired characteristics of information assurance approaches for particular problems. While this Framework attempts to lay out a large amount of information in an orderly sequence, it is structured to allow readers to use the table of contents to find topics of interest.

# Appendix A

## Acronyms

AAA	Authentication, Authorization and Accounting
ACDF	Access Control Decision Function
ACI	Access Control Information
ACL	Access Control List
ACN	Airborne Communications Node
ADRP	Army's DISN Router Program
ADSL	Asymmetric Digital Subscriber Line
AFIWC	Air Force Information Warfare Center
AH	Authentication Header
AIS	Automated Information System
AJ	Anti-Jam
AMPS	Advanced Mobile Phone Service
AMSC	American Mobile Satellite Corporation
ANX	Automotive Network eXchange®
API	Application Program Interface
ASCII	American Standard Code for Information Interchange
ASD(C³I)	Assistant Secretary of Defense (Command, Control, Communications, and Intelligence)
ASN1	Abstract Syntax Notation
ATF	Alcohol, Tobacco, and Firearms
ATM	Asynchronous Transfer Mode
AV	Anti-Virus
BAPI	Biometrics Application Program Interface
BCA	Bridge Certificate Authority
BFACS	Battlefield Functional Area Control System



**UNCLASSIFIED**

Appendix A  
IATF Release 3.0—September 2000

BIOS	Basic Input/Output System
BIOS ROM	Basic Input/Output System Read Only Memory
BN	Backbone Network
BOOTP	Boot Protocol
BSD	Berkeley Software Design
BST	Base Support Teams
BVTC	Battlefield Video Teleconferencing
C&A	Certification and Accreditation
C/AII	Corporate/Agency Information Infrastructure
C2	Command and Control
C4I	Command, Control, Communications, Computer and Intelligence
CA	Certification Authority
CALEA	Communications Assistance for Law Enforcement Act
CAN	Campus Area Network
CAPI	Cryptographic Application Programming Interface
CAT	Common Authentication Technology
CAW	Certificate Authority Workstation
CB	Citizens Band
CBR	Constant Bit Rate Case-Based Reasoning
CC	Common Criteria
CCI	Controlled Cryptographic Item
CCITT	Consultative Committee for International Telephone and Telegraph
CD	Compact Disc
CDMA	Code Division Multiple Access
CDR	Critical Design Review
CD-ROM	Compact Disc-Read Only Memory
CDSA	Common Data Security Architecture
CECOM	Communications Electronics Command
CEOI	Communications-Electronics Operating Instruction
CERT	Computer Emergency Response Team

CFD	Common Fill Devices
CGI	Common Gateway Interface
CH	Correspondent Host
CI	Cryptographic Interface Configuration Item
CIAC	Computer Incident Advisory Capability
CIDF	Common Intrusion Detection Framework
CIK	Crypto-Ignition Key
CINCS	Commander In Chiefs
CIO	Chief Information Officer
CITS	Combat Information Transport System
CKL	Compromised Key List
CM	Configuration Management
CMA	Certificate Management Authority
CMI	Certificate Management Infrastructure
CMIP	Common Management Information Protocol
CMP	Certificate Management Protocols
CMS	Certificate Management System
CMUA	Certificate Management User Agent
CMW	Compartmented Mode Workstations
CND	Computer Network Defense
CNR	Combat Net Radio
CO	Central Office
COA	Course Of Action
COE	Common Operating Environment
COI	Community of Interest
COMSEC	Communications Security
CONOPS	Concept of Operations
CONUS	Continental United States
CORBA	Common Object Request Broker Architecture
COTS	Commercial-Off-The-Shelf
CP	Certificate Policy

**UNCLASSIFIED**

Appendix A  
IATF Release 3.0—September 2000

CPS	Certification Practice Statement
CPU	Central Process Unit
CRC	Cyclic Redundancy Check
CRL	Certificate Revocation List or Compromise Recovery List
CryptoAPI	Cryptographic Application Program Interface
CSN	Central Services Node
CSP	Cryptographic Service Provider
CSRA	Critical Security Requirement Areas
CSSM	Common Security Services Manager
CTP	Common Tactical Picture
CUG	Closed User Group
CUITN	Common User Installation Transport Network
CV	Compliance Validation
CVI	Compliance Validation Inspection
CVSD	Continuously Variable Slope Detection
DA	Directory Administrators
DAA	Designated Approving Authority
DAC	Discretionary Access Control
DAP	Directory Access Protocol
DARPA	Defense Advanced Research Projects Agency
DCOM	Distributed Component Object Model
DEA	Drug Enforcement Agency
DECT	Digital Enhanced Cordless Telecommunications
DEERS	Defense Eligibility and Enrollment Reporting System
DER	Distinguished Encoding Rules
DES	Data Encryption Standard
DFAS	Defense Finance and Accounting Service
DHCP	Dynamic Host Control Protocol
DIAP	Defense Information Assurance Program
DIB	Directory Information Base
DII	Defense Information Infrastructure

DISA	Defense Information Systems Agency
DISN	Defense Information Systems Network
DIT	Directory Information Tree
DITSCAP	Department of Defense Information Technology Security Certification and Accreditation Process
DMS	Defense Message System
DMZ	Demilitarized Zone
DN	Distinguished Name Defined Name
DNA	Deoxyribose Nucleic Acid
DNS	Domain Name Server
DNSSEC	Domain Name System Security
DoD	Department of Defense
DOE	Department of Energy
DOS	Denial of Service
DSA	Directory Server Agents
DSL	Digital Subscriber Line
DSN	Defense Switched Network
DSP	Digital Signal Processing
DSS	Digital Signature Standard
DSSMP	Digital Switched Systems Modernization Program
DSSS	Direct Sequence Spread Spectrum
DTD	Data Transfer Device
E911	Emergency 911
EAL	Evaluation Assurance Level
ECAs	External Certificate Authorities
ECP	Engineering change proposal
EEPROM	Electrically Erasable Programmable Read Only Memory
EISL	Embedded Integrity Services Library
EKMS	Electronic Key Management System
E-mail	Electronic Mail

**UNCLASSIFIED**

Appendix A  
IATF Release 3.0—September 2000

EPLRS	Enhanced Position/Location Reporting System
EPROM	Erasable Programmable Read Only Memory
ESM	Enterprise Security Management
ESNet	Department of Energy Science Network
ESP	Encapsulating Security Payload
EUT	End User Terminal
FAA	Federal Aviation Administration
FBI	Federal Bureau of Investigation
FCC	Federal Communications Commission
FDMA	Frequency Division Multiple Access
FEDCERT	Federal Computer Emergency Response Team
FEMA	Federal Emergency Management Agency
FFC	Fortezza for Classified
FHSS	Frequency Hopped Spread Spectrum
FIDNet	Federal Intrusion Detection Network
FIPS	Federal Information Processing Standard
FIRST	Forum of Incident Response and Security Team
FNBDT	Future Narrow Band Digital Terminal
FPKI	Federal Public Key Infrastructure
FRP	Federal Response Plan
FSRS	Functional Security Requirements for Specification
FTP	File Transfer Protocol
FW	Firewall
GAO	Government Accounting Office
GBS	Global Broadcast Service
GCCS	Global Command and Control System
GIG	Global Information Grid
GII	Global Information Infrastructure
GNIE	Global Network Information Enterprise
GNOSC	Global Network Operations and Security Center

GOTS	Government-Off-The-Shelf
GPS	Global Positioning System
GSA	General Services Administration
GSA KMP	Group Service Association Key Management Protocol
GSM	Groupe Speciale Mobile (now known as the Global System for Mobile Communications)
GSS-API	Generic Security Services Application Programming Interface
GUI	Graphical User Interface
GULS	General Upper Layer Security
HAG	High Assurance Guard
HF	High Frequency
HTML	HyperText Markup Language
HTTP	HyperText Transfer Protocol
I&A	Identification and Authentication
IA	Information Assurance
IATF	Information Assurance Technical Framework
IBAC	Identity Based Access Control
IC	Intelligence Community
ICMP	Internet Control Message Protocol
ICRLA	Indirect Certificate Revocation List Authority
ICSA	International Computer Security Association
ID	Identification
IDS	Intrusion Detection System
IDUP	Independent Data Unit Protection
IEC	International Engineering Consortium
IEEE	Institute of Electrical and Electronics Engineers
IETF	Internet Engineering Task Force
IKE	Internet Key Exchange
ILS	Integrated Logistics Support
IMAP	Internet Mail Access Protocol

**UNCLASSIFIED**

Appendix A  
IATF Release 3.0—September 2000

IMMARSAT	Internal Maritime Satellite
INE	In-line Network Encryptor
INFOCON	Information Condition
INFOSEC	Information Security
INS	Immigration and Naturalization Service
IOS	Internet Operating System
IP	Internet Protocol
IPN	Information Protection Network
IPSec	Internet Protocol Security
IPX	Internet Packet eXchange
IR	Infrared
IS	Information Systems
ISAC	Information Sharing and Analysis Center
ISAKMP	Internet Security Association and Key Management Protocol
ISDN	Integrated Services Digital Network
ISO	International Organization for Standardization
ISP	Internet Service Provider
ISSE	Information System Security Engineers Information System Security Engineering
ISSO	Information Systems Security Organization
IT	Information Technology
IT-21	Information Technology for the 21 <sup>st</sup> Century
ITU	International Telecommunications Union
IW	Information Warfare
JNMS	Joint Network Management System
JPO	Joint Program Office
JTF-CND	Joint Task Force for Computer Network Defense
JTIDS	Joint Tactical Information Distribution System
JTRS	Joint Tactical Radio System
JWICS	Joint Worldwide Intelligence Communications System

KMI	Key Management Infrastructure
KP	Key Processor
KRA	Key Recovery Agent
KRB	Key Recovery Block
KRF	Key Recovery Field
KRI	Key Recovery Information
KU	KMI Unique
LAN	Local Area Network
LCC	Life-Cycle Costing
LDAP	Lightweight Directory Access Protocol
LEO	Low Earth Orbit
LMD/KP	Local Management Device/Key Processor
LMR	Land Mobile Radio
LOS	Line-Of-Site
LPC/CELP	Linear Predictive Coding/Codebook Excited Linear Prediction
LPD	Low Probability of Detection
LPI	Low Probability of Intercept
LRA	Local Registration Authority
LSE	Local Subscriber Environment
MAC	Mandatory Access Control or Media Access Control
MAN	Metropolitan Area Network
MANET	Mobile Ad Hoc Networking
MATTS	Mobile Air Transportable Telecommunications System
MBR	Model-Based Reasoning
MCS	Maneuver Control System
MD4	Message Digest 4
MEO	Medium Earth Orbit
MERS	Mobile Emergency Response Support
MIB	Management Information Base



**UNCLASSIFIED**

Appendix A  
IATF Release 3.0—September 2000

MIL STD	Military Standard
MILDEP	Military Department
MILS	Multiple, Independent Security Levels
MILSATCOM	Military Satellite Communications
MIME	Multipurpose Internet Mail Extension
MISSI	Multi-Level Information Systems Security Initiative
MIT	Massachusetts Institute of Technology
MLS	Multi-Level Security
MMP	MISSI Management Protocol
MNS	Mission Needs Statement
MOB	Main Operating Base
MoE	Measure of Effectiveness
MPP	Moving Pictures Expert Group
MSE	Mobile Subscriber Equipment
MSP	Message Security Protocol
MSS	Mobile Satellite Subscriber Mobile Satellite Service
MTA	Message Transfer Agent
MTBF	Mean Time Between Failure
MTS	Mail Transfer System
MTSC	Mobile Telephone Switching Center
MTTR	Mean Time to Repair
NAT	Network Address Translation
NATO	North Atlantic Treaty Organization
NAVCERT	Navy Computer Emergency Response Team
NBC	Nuclear, Biological, and Chemical
NCO	Non-Commissioned Officer
NES	Network Encryption System
NETBIOS	Network Basic Input/Output System
NIAP	National Information Assurance Partnership
NIC	Network Interface Card

NII	National Information Infrastructure
NIPRNET	Non-classified Internet Protocol Router Network
N-ISDN	Narrowband Integrated Services Digital Network
NIST	National Institute of Standards and Technology
NMC	Network Management Center
NNTP	Network News Transfer Protocol
NORAD	North American Aerospace Defense
NOS	Network Operating Systems
NS/EP	National Security/Emergency Preparedness
NSA	National Security Agency
NSF	Network Security Framework
NSIRC	National Security Incident Response Center
NSTISSP	National Security Telecommunications and Information Systems Security Policy
OA&M	Operations, Administration and Maintenance
OAN	Operational Area Network
OASD(AS (C3I))	Office of the Assistant Secretary of Defense (Command, Control, Communications, and Intelligence)
OPSEC	Operational Security
ORD	Operational Requirements Document
OS	Operating System
OSCAR	Outside Cable Rehabilitation
OSD	Office of the Secretary of Defense
OSI	Open Systems Interconnection
OTAR	Over-the-Air Rekey
OTAT	Over-the-Air Transfer
OTAZ	Over-the-Air Zeroize
PAA	Policy Approving Authority
PBX	Private Branch eXchange
PC	Personal Computer

**UNCLASSIFIED**

Appendix A  
IATF Release 3.0—September 2000

PCA	Policy Creation Authority
PCCIP	President's Commission on Critical Infrastructure Protection
PCI	Protocol Control Information
PCM	Pulse Code Modulation
PCMCIA	Personal Computer Memory Card International Association
PCS	Personal Communications System
PDA	Personal Digital Assistant
PDD	Presidential Decision Directive
PDR	Preliminary Design Review
PEIP	Programmable Embedded INFOSEC Program
PERL	Practical Extraction and Report Language
PGP	Pretty Good Privacy
PHS	Personal Handyphone System
PIN	Personal Identification Number
PK	Public Key
PKCS	Public Key Cryptographic Standards
PKI	Public Key Infrastructure
PM	Privilege Manage
PMA	Policy Management Authority
PMO	Program Management Office
PNA	Protection for Network Access
POP	Post Office Protocol Proof of Possession
POSIX	Portable Operating System Interface Unix
POTS	Plain Old Telephone Service
PP	Protection Profile
PPP	Point-to-Point Protocol
PROM	Programmable Read Only Memory
PRSN	Primary Services Node
PSN	Production Source Node
PSTN	Public Switched Telephone Network
PVC	Permanent Virtual Circuit

QOP	Quality of Protection
QOS	Quality of Service
R&D	Research and Development
RA	Registration Authorities
RADIUS	Remote Access Dial In User Service
RAM	Random Access Memory
RBAC	Rule Based Access Control
RBR	Rule-Based Reasoning
RDN	Relative Distinguished Names
RF	Radio Frequency
RFC	Request for Comment
RIM	Recovery Information Medium
RM	Registration Managers
RMON	Remote Monitor
RSVP	Resource Reservation Protocol
RTM	Requirements Traceability Matrix
S/MIME	Secure/Multipurpose Internet Mail Extension
SA	System Administrator
SABI	Secret and Below Initiative
SATCOM	Satellite Communications
SBU	Sensitive But Unclassified
SC	Steering Committee
SCIF	Secure Compartmented Information Facility
SDD	Secure Data Device
SDE	Secure Data Exchange
SDLS	Single-Line Digital Subscriber Line
SDR	Software Designed Radio
SE	Systems Engineering

**UNCLASSIFIED**

Appendix A  
IATF Release 3.0—September 2000

SET	Secure Electronic Transaction
S-FTP	Secure File Transfer Protocol
SFUG	Security Features Users Guide
S-HTTP	Secure HyperText Transfer Protocol
SID	System Identification
SIM	Subscriber Identity Module
SINCGARS	Single Channel Ground and Airborne Radio System
SIPRNET	Secret Internet Protocol Router Network
SKM	Symmetric Key Management
SLA	Service Level Agreements
SLIP	Serial Line Internet Protocol
SMB	Server Message Block
SMDS	Switched Multi-megabit Data Service
SMI	Security Management Infrastructure
SMIB	Security Management Information Base
SML	Strength of Mechanism Level
SMTP	Simple Mail Transfer Protocol
SNMP	Simple Network Management Protocol
SONET	Synchronous Optical NETwork
SPG	Security Program Group
SPI	Service Provider Interface
SPI_NET	Security Profile Inspector For Network
SPKI	Simple Public Key Infrastructure
SS7	Signaling System 7
SSA	System Security Administrator
SSAA	System Security Authorization Agreement
SSAPI	Microsoft's Security Service Provider
SSE CMM	System Security Engineering Capability Maturity Model
SSH	Secure Shell
SSL	Secure Sockets Layer
ST	Security Target
STAMIS	Standard Army Management Information System

STE	Secure Telephone Equipment
STG	State Transition Graphs
STU	Secure Telephone Unit
STS	Synchronous Transport Service
SVC	Switched Virtual Circuit
TAC/STE	Tactical Secure Telephone Equipment
TCB	Trusted Computing Base
TCP	Transmission Control Protocol
TCP/IP	Transmission Control Protocol/Internet Protocol
TDC-ICAP	Theater Deployable Communications Integrated Communications Access Package
TDMA	Time Division Multiple Access
TDY	Temporary Duty
TISM	TACLANE Internet Security Manager
TLS	Transport Layer Security
TOC	Tactical Operations Center
TOE	Target of Evaluation
TPEP	Trusted Product Evaluation Program
TPN	Tactical Packet Network
TRANSCOM	Transportation Command
TRANSEC	Transmission Security
TRI-TAC	Tri-Tactical
TS	Top Secret
TSDM	Trusted Software Design
TS-SCI	Top Secret-Sensitive Compartmented Information
TTP	Trusted Third Party
TWG	Technical Working Group
U	Unclassified
U.S.	United States
UAV	Unmanned Aerial Vehicle

**UNCLASSIFIED**

Appendix A  
IATF Release 3.0—September 2000

UBS	Unclassified But Sensitive
UDP	User Datagram Protocol
UHF	Ultra-High Frequency
UPS	Uninterruptible Power Supply
URL	Universal Resource Locator
VA	Veterans Affairs
VBA	Visual BASIC Application
VDSL	Very High Bit-Rate Digital Subscriber Line
VHF	Very-High Frequency
VM	Virtual Machine
VPN	Virtual Private Network
VRML	Virtual Reality Modeling Language
VSAT	Very Small Aperture Terminal
VTC	Video Conferencing
WAIS	Wide-Area Information Service
WAN	Wide Area Network
WIN	Warfighter Information Network or Wireless Intelligent Network
WLAN	Wireless Local Area Network
WLL	Wireless Local Loop
WWW	World Wide Web

# Appendix B

## Glossary

Access Control	A security service that prevents the unauthorized use of information system resources (hardware and software) only to authorized users and the unauthorized disclosure or modification of data (stored and communicated).
Access Control List (ACL)	Mechanism implementing discretionary and/or mandatory access control between subjects and objects.
Accountability	Property that allows auditing of activities in an automated information system (AIS) to be traced to persons who may then be held responsible for their actions.
Advanced Mobile Phone Service (AMPS)	The standard system for analog cellular telephone service in the U.S. AMPS allocates frequency ranges within the 800 – 900 MHz spectrum to cellular telephones. Signals cover an area called a cell. Signals are passed into adjacent cells as the user moves to another cell. The analog service of AMPS has been updated to include digital service.
Adversary	Person or organization that must be denied accesses to information.
Anonymity	A security service that prevents the disclosure of information that leads to the identification of the end users.
Anti-Jam	Measures ensuring that transmitted information can be received despite deliberate jamming attempts.
Application-Level Firewall	A firewall system in which service is provided by processes that maintain complete TCP connection state and sequencing; application level firewalls often re-address traffic so that outgoing traffic appears to have originated from the firewall, rather than the internal host. In contrast to packet filtering firewalls, this firewall must have knowledge of the application data transfer protocol and often has rules about what may be transmitted and what may not.
Application Program Interface (API)	A set of standard software interrupts, calls, and data formats that application programs use to initiate contact with network services, main-frame communications programs, telephone equipment, or program-to-program communications.



**UNCLASSIFIED**

Appendix B  
IATF Release 3.0—September 2000

Assurance	A measure of confidence that a security feature and architecture of an automated information system mediates and enforces a security policy.
Asymmetric Cryptographic Algorithm	An encryption algorithm that requires two different keys for encryption and decryption. These keys are commonly referred to as the public and private keys. Asymmetric algorithms are slower than symmetric algorithms. Furthermore, speed of encryption may be different than the speed of decryption. Generally asymmetric algorithms are either used to exchange symmetric session keys or to digitally sign a message. RSA, RPK, and ECC are examples of asymmetric algorithms.
Asynchronous Transfer Mode (ATM)	A dedicated connection switching technology that organizes digital data into fixed byte cell units and transmits those units over a physical medium using digital signal technology. It is implemented by hardware, therefore, very fast processing and switching speeds are possible.
Attack	The intentional act of attempting to bypass security controls on an automated information system.
Audit	A service that keeps a detailed record of events.
Authentication	Security measure designed to establish the validity of a transmission, message or originator, or a means of verifying an individual's eligibility to receive specific categories of information.
Authentication Header (AH)	An IP device used to provide connectionless integrity and data origin authentication for IP datagrams.
Authentication Token	See token.
Authorization	The process of determining what types of activities is permitted; Usually, authorization is in the context of authentication: once you have authenticated a user, they may be authorized different types of access or activity.
Automated Information System (AIS)	The entire infrastructure, organization, personnel, and components for the collection, processing, storage, transmission, display, dissemination, and disposition of information.
Availability	The property of being accessible and usable upon demand by an authorized entity.

Biometrics	Automated methods of authenticating or verifying an individual based on a physical or behavioral characteristic.
CERT	Computer Emergency Response Team – A federally funded research and development center at Carnegie Mellon University. They focus on Internet security vulnerabilities, provide incident response services to sites that have been the victims of attack, publish security alerts, research security and survivability in wide-area-networked computing, and develop site security information. They can be found at <a href="http://www.cert.org">www.cert.org</a> .
Certificate	Record holding security information about an AIS user and vouches to the truth and accuracy of the information it contains.
Certificate Revocation List (CRL)	A record of all revoked certificates produced by a common Issuer; a certificate is revoked when any data in it changes before it expires, e.g., when a user moves and changes addresses.
Certification and Accreditation (C&A)	Certification is the comprehensive evaluation of the technical and non-technical security features of an IS and other safeguards, made in support of the accreditation process, to establish the extent to which a particular design and implementation meets a set of specified requirements. Accreditation is the formal declaration by a DAA that an IS approved to operate in a particular security mode using a prescribed set of safeguards at an acceptable level of risk.
Certification Authority (CA)	A trusted agent that issues digital certificates to principals. Certification authorities may themselves have a certificate that is issued to them by other certification authorities. The highest certification authority is called the root CA.
Challenge/Response	An authentication technique whereby a server sends an unpredictable challenge to the user, who computes a response using some form of authentication token.
Checksum	Value computed, via some parity or hashing algorithm, on information requiring protection against error or manipulation.
Code Division Multiple Access (CDMA)	A digital cellular phone spread spectrum technology that assigns a code to all speech bits, sends a scrambled transmission of the encoded speech over the air and reassembles the speech to its original format.
Common Criteria (CC)	The Common Criteria for Information Technology Security Evaluation [CC98] is a catalog of security functional and assurance requirements and has a central role in the National Information Assurance Program.

**UNCLASSIFIED**

Appendix B  
IATF Release 3.0—September 2000

Communications Security	Measures and controls taken to deny unauthorized persons information derived from telecommunications and to ensure the authenticity of such telecommunications.
Compromise	Disclosure of information of data to unauthorized persons, or a violation of the security policy of a system in which unauthorized intentional or unintentional disclosure, modification, destruction, or loss of an object may have occurred.
Compromised Key List (CKL)	A list with the Key Material Identifier (KMID) of every user with compromised key material; key material is compromised when a card and its personal identification number (PIN) are uncontrolled or the user has become a threat to the security of the system.
Computer Intrusion	An incident of unauthorized access to data or an Automated Information System (AIS).
Computer Security	Measures and controls that ensure confidentiality, integrity and availability of information system assets including hardware, software, firmware and information being processed, stored, or communicated.
CONcept of Operations (CONOP)	Describes how the system would be used to accomplish objectives.
Confidentiality	A security service that prevents unauthorized disclosure of information residing on a computer, transiting a local network, or flowing over a public Internet.
Configuration Management (CM)	Management of security features and assurances through control of changes made to hardware, software, firmware, documentation, test, test fixtures, and test documentation throughout the life-cycle of an IS.
Countermeasure	A specific technique, product or procedure that is implemented to subvert or remedy the effects of an attack or attack scenario.
Covert Channel	Any communication channel that can be exploited by a process to transfer information in a manner that violates the system's security policy.
Credential	A credential is the information one entity presents to another to authenticate the other's identity.
Crypto-Ignition Key (CIK)	Device or electronic key used to unlock the secure mode of crypto-equipment.

Cryptographic Application Program Interface	An interface standard that provides a means for isolating a computer platform from the details of the implementation of cryptographic functions.
Cryptographic Function	A set of procedures that provide basic cryptographic functionality using various algorithms for key generation, random number generation, encryption, decryption, and message digesting.
Customer	See User.
Data Encryption Standard (DES)	A 56-bit, private key, symmetric cryptographic algorithm for the protection of unclassified computer data issued as Federal Information Processing Standard Publication.
Defense in Depth	The security approach whereby layers of protection are needed to establish an adequate security posture for a system; strategy is based on concept that attacks must penetrate multiple protections that have been placed throughout the system to be successful.
Defense-wide Information Assurance Program (DIAP)	This Department of Defense (DoD) program provides for the planning, coordination, integration, and oversight of the DoD information assurance resources to assure the availability, integrity, authentication, confidentiality, and non-repudiation of the DoD's mission essential and mission support information.
Denial of Service (DOS)	An attack where service is denied to legitimate users.
Designated Approval Authority (DAA)	The DAA determines the level of acceptable risk for a system and authorizes the operation of an information system by issuing an accreditation statement once an acceptable level of risk has been obtained.
Digital Certificate	A structure for binding a principal's identity to its public key. A certification authority (CA) issues and digitally signs a digital certificate.
Digital Signature	A method for verifying that a message originated from a principal and that it has not changed en route. Digital signatures are typically generated by encrypting a digest of the message with the private key of the signing party.
Discretionary Access Control (DAC)	A non-policy-based method of restricting access to a system's files/objects based on the decision of the resource's owner.

**UNCLASSIFIED**

Appendix B  
IATF Release 3.0—September 2000

DoD Information Technology Security Certification and Accreditation Process (DITSCAP)	The standard DoD approach for identifying information security requirements, providing security solutions, and managing information technology system security. (DoDI 5200.40)
Downgrade	The change of a classification label to a lower level without changing the contents of the data. Downgrading occurs only if the content of a file meets the requirements of the sensitivity level of the network for which the data is being delivered.
Eavesdropping	An attack in which an attacker listens to a private communication. The best way to thwart this attack is by making it very difficult for the attacker to make any sense of the communication by encrypting all messages.
Effective Key Length	A measure of strength of a cryptographic algorithm, regardless of actual key length.
Electronic Key Management System (EKMS)	The EKMS is a National Security Agency (NSA) effort to electronically provide communications security (COMSEC) material and provide a logistics support system consisting of interoperable Department, Agency or Organization (DAO) key management systems.
Encapsulating Security Payload	This message header is designed of provide a mix of security services that provides confidentiality, data origin authentication, connectionless integrity, an anti-replay service, ad limited traffic flow confidentiality.
Evaluation Assurance Level (EAL)	A scale for measuring the criteria for the evaluation of requirements. Uniformly increasing, the scale balances the level of assurance obtained with the cost and feasibility of acquiring that level of assurance.
Firewall	A system or combination of systems that enforces a boundary between two or more networks.
Frequency Division Multiple Access (FDMA)	A technique for sharing a single transmission channel, such as a satellite transponder, among two or more users by assigning each to an exclusive frequency band within the channel.
Future Narrow Band Digital Terminal (FNBDT)	It is a network-independent/transport-independent message layer. FNBDT operates in the Narrow Band portion of the STE spectrum (64 kbps and below).

Global Information Grid (GIG)	It is a globally interconnected, end-to-end set of information capabilities, associated processes and personnel for collecting, processing, storing, disseminating, and managing information on demand to warfighters, policy makers, and support personnel.
Global Command and Control system (GCCS)	A comprehensive, worldwide network of systems that provide the NCA, Joint staff, combatant and functional unified commands, services, and defense agencies, Joint Task Forces and their service components, and others with information processing and dissemination capabilities necessary to conduct C2 of forces.
Global Network Information Environment (GNIE)	A composition of all information system technologies used to process, transmit, store, or display DoD information. It has been superseded by Global Information Grid (GIG).
Guard	A highly assured device that negotiates the transfer of data between enclaves operating at different security levels.
Hacker	Unauthorized user who attempts or gains access to an AIS.
Hash	Value computed on data to detect error or manipulation. See checksum.
Host-based Security	The technique of securing an individual system from attack; host-based security is operating system and version dependent.
Identification	Process that enables recognition of an entity by an AIS.
Identification & Authentication (I&A)	Identity of an entity with some level of assurance.
Information Assurance (IA)	Information operations (IO) that protect and defend information and information systems by ensuring their availability, integrity, authentication, confidentiality, and non-repudiation. This includes providing for restoration of information systems by incorporating protection, detection, and reaction capabilities.
Information Protection Policy	See Security Policy.

**UNCLASSIFIED**

Appendix B  
IATF Release 3.0—September 2000

Information Systems Security (INFOSEC)	Protection of information systems against unauthorized access to or modification of information, whether in storage, processing or transit, and against the denial of service to authorized users, including those measures necessary to detect, document, and counter such threats.
Information Systems Security Engineering (ISSE)	A structured system engineering process, tailored to the unique needs of a specific customer, focused on the selection of an effective security protection solution including identifying the customer's requirements, determining the a-priori/initial vulnerabilities and threats of the existing or planned communications system, determining applicable security solutions and/or countermeasures, identifying the residual risk, and implementing a risk management process to determine if the risk is acceptable to the customer.
Information Technology (IT)	The hardware, firmware, and software used as part of the information system to perform DoD information functions. This definition includes computers, telecommunications, automated information systems, and automatic data processing equipment as well as any assembly of computer hardware, software, and/or firmware configured to collect, create, communicate, compute, disseminate, process, store and/or control data or information.
Insider Attack	An attack originating from inside a protected network.
Integrity [of data]	A security service that allows verification that an unauthorized modification of information (including changes, insertions, deletions, and duplications) has not occurred either maliciously or accidentally.
Internet Control Message Protocol – ICMP	A message control and error-reporting protocol between a host server and a gateway to the Internet. ICMP is used by a device, often a router, to report and acquire a wide range of communications-related information.
Intrusion Detection	Detection of break-ins or break-in attempts either manually or via software expert systems that operate on logs or other information available on the network.
Intrusion Detection System (IDS)	A system that detects and identifies unauthorized or unusual activity on the hosts and networks; this is accomplished by the creation of audit records and checking the audit log against the intrusion thresholds.
Key	Usually a sequence of random or pseudo-random bits used initially to set up and periodically change the operations performed in crypto-equipment for the purpose of encrypting or decrypting electronic signals, or for producing other keys.

Key Management	Supervision and control of the process whereby key is generated, stored, protected, transferred, loaded, used, and destroyed.
Key Management Infrastructure (KMI)	Framework established to issue, maintain, and revoke keys accommodating a variety of security technologies, including the use of software.
Labeling	Process of assigning a representation of the sensitivity of a subject or object
Layered Solution	The judicious placement of security protections and attack countermeasures that can provide an effective set of safeguards that are tailored to the unique needs of a customer's situation.
Local Area Network (LAN)	A limited distance, high-speed data communication system that links computers into a shared system (two to thousands) and is entirely owned by the user. Cabling typically connects these networks.
Mandatory Access Control (MAC)	Policy-based control methods of restricting access to a system's file/objects in which the administrators, not the resource owners, make access decisions that bear on or derive from access control policy.
Mission Needs Statement (MNS)	Describes the mission need or deficiency; identifies threat and projected threat environment
Motivation	The specific technical goal that a potential adversary wants to achieve by an attack, e.g., gain unauthorized access, modify, destroy or prevent authorized access.
Multi Level Security	A system that can simultaneously process data communications at different levels of classification while enforcing secure access and authorization.
Multipurpose Internet Mail Extensions (MIME)	A specification for formatting non-ASCII messages so that they can be sent over the Internet. MIME enables graphics, audio, and video files to be sent and received via the Internet mail system. In addition to e-mail applications, Web browsers also support various MIME types. This enables the browser to display or output files that are not in HTML format. The Internet Engineering Task Force (IETF) defined MIME in 1992. See also Secure Multipurpose Internet Mail Extensions, S/MIME.
National Information Assurance Partnership (NIAP)	A joint industry/government initiative, lead by NIST and NSA, to establish commercial testing laboratories where industry product providers can have security products tested to verify their performance against vendor claims.



**UNCLASSIFIED**

Appendix B  
IATF Release 3.0—September 2000

Non-Repudiation	A security service by which evidence is maintained so that the sender of data and recipient of data cannot deny having participated in the communication.
Non-Technical Countermeasure	A security measure, that is not directly part of the network information security processing system, taken to help prevent system vulnerabilities. Non-technical countermeasures encompass a broad range of personnel measures, procedures, and physical facilities that can deter an adversary from exploiting a system.
Open System Interconnection Model (OSI)	A reference model of how messages should be transmitted between any two endpoints of a telecommunication network. The process of communication is divided into seven layers, with each layer adding its own set of special, related functions. The seven layers are the application layer, presentation, session, transport, network, data, and physical layer. Most telecommunication products tend to describe themselves in relation to the OSI model. The OSI model is a single reference view of communication that provides a common ground for education and discussion.
Parity	Bit(s) used to determine whether a block of data has been altered.
Password	Protected/private character string used to authenticate an entity or to authorize access to data.
Perimeter-based Security	The technique of securing a network by controlling accesses to all entry and exit points of the network.
Pretty Good Privacy (PGP)	A standard program for securing e-mail and file encryption on the Internet. Its public-key cryptography system allows for the secure transmission of messages and guarantees authenticity by adding digital signatures to messages.
Protection Needs Elicitation (PNE)	Discovering the customer's prioritized requirements for the protection of information.
Protection Profile (PP)	An implementation-independent set of security requirements for a category of TOEs that meet specific consumer needs.

Proxy	A software agent that acts on behalf of a user. Typical proxies accept a connection from a user, make a decision as to whether or not the user or client IP address is permitted to use the proxy, perhaps does additional authentication, and then completes a connection on behalf of the user to a remote destination.
Public Key Infrastructure (PKI)	Public and private keys, digital certificates, certification authorities, certificate revocation lists, and the standards that govern the use and validity of these elements make up an infrastructure where principals can engage in private and non-repudiable transactions. This combination is called the Public Key Infrastructure (PKI).
Residual Risk	The risk remaining in an information system or network after the implementation of security countermeasures.
Risk	A situation where there is a known vulnerability and a potential adversary with the motivation and capability to exploit that vulnerability.
Risk Plane	A graphic technique for depicting the likelihood of particular attacks occurring and the degree of consequence to an operational mission.
Robustness	A characterization of the strength of a security function, mechanism, service, or solution, and the assurance (or confidence) that is implemented and functioning correctly.
Sanitization –	The changing of content information in order to meet the requirements of the sensitivity level of the network to which the information is being sent.
Secret Key	A key used by a symmetric algorithm to encrypt and decrypt data.
Secure Multipurpose Internet Mail Extensions - S/MIME	A version of the MIME protocol that supports encrypted messages. S/MIME is based on RSA's public-key encryption technology. See also Multipurpose Internet Mail Extensions, MIME.
Secure State	Condition in which no subject can access or utilize any object in an unauthorized manner.
Security Management Infrastructure (SMI)	A set of interrelated activities providing security services needed by other security features and mechanisms; SMI functions include registration, ordering, key generation, certificate generation, distribution, accounting, compromise recovery, rekey, destruction, data recovery, and administration.

## UNCLASSIFIED

Appendix B  
IATF Release 3.0—September 2000

Security Policy	What security means to the user; a statement of what is meant when claims of security are made. More formally, it is the set of rules and conditions governing the access and use of information. Typically, a security policy will refer to the conventional security services, such as confidentiality, integrity, availability, etc., and perhaps their underlying mechanisms and functions.
Security Target (ST)	A set of security requirements and specifications to be used as the basis for evaluation of an identified TOE.
Session Key	A temporary symmetric key that is only valid for a short period. Session keys are typically random numbers that can be chosen by either party to a conversation, by both parties in cooperation with one another, or by a trusted third party.
Signature [Digital, Electronic]	A process that operates on a message to assure message source authenticity and integrity, and may be required for source non-repudiation.
Social Engineering	An attack based on deceiving users or administrators at the target Site and are typically carried out by an adversary telephoning users or operators and pretending to be an authorized user, to attempt to gain illicit access to systems.
SOCKS	A networking proxy protocol that enables full access across the SOCKS server from one host to another without requiring direct IP reachability. The SOCKS server authenticates and authorizes the requests, establishes a proxy connection, and transmits the data. SOCKS is commonly used as a network firewall that enables hosts behind a SOCKS server to gain full access to the Internet, while preventing unauthorized access from the Internet to the internal hosts.
Strength of Mechanism (SML)	A scale for measuring the relative strength of a security mechanism hierarchically ordered from SML 1 through SML 3.
Symmetric Algorithm	An algorithm where the same key can be used for encryption and decryption.
System Security Authorization Agreement (SSAA)	The SSAA is the formal agreement among the DAA(s), Certifier, user representative, and program manager. It is used throughout the entire DITSCAP to guide actions, document decisions, specify IA requirements, document certification tailoring and level-of-effort, identify potential solutions, and maintain operational systems security.

Tamper	Unauthorized modification that alters the proper functioning of cryptographic or automated information system security equipment in a manner that degrades the security or functionality it provides.
Target of Evaluation (TOE)	An IT product or system and its associated administrator and user guidance documentation that is the subject of an evaluation.
Technical Countermeasure	A security feature implemented in hardware and/or software, that is incorporated in the network information security processing system.
Technology Gap	A technology that is needed to mitigate a threat at a sufficient level but is not available.
TEMPEST	The investigation, study and control of compromising emanations from telecommunications and automated information systems equipment.
Threat	Capabilities, intentions, and attack methods of adversaries to exploit, or any circumstance or event with the potential to cause harm to, information or an information system.
Time Division Multiple Access (TDMA)	A technique to interweave multiple conversations into one transponder so as to appear to get simultaneous conversations.
Token	A token is used to validate an end entity's identity and bind that identity to its public key. An example is an X.509 certificate.
Trojan Horse	A software entity that appears to do something normal but which in fact contains a trapdoor or attack program.
Trusted Computing Base (TCB)	The totality of protection mechanisms within a computer system, the combination of which is responsible for enforcing a security policy.
Trusted Operating System	An operating system that satisfies a number of stringent security requirements where high security is required.
Tunneling Router	A router or system capable of routing traffic by encrypting it and encapsulating it for transmission across an untrusted network, for eventual de-encapsulation and decryption.
User	The party, or his designee, responsible for the security of designated information. The user works closely with an ISSE. Also referred to as the customer.

**UNCLASSIFIED**

Appendix B  
IATF Release 3.0—September 2000

Virtual Network Perimeter	A network that appears to be a single protected network behind firewalls, which actually encompasses encrypted virtual links over untrusted networks.
Virtual Private Network (VPN)	Virtual Private Network; a way of using a public network (typically the Internet) to link two sites of an organization. A VPN is typically set up by protecting the privacy and integrity of the communication line using a secret session key. The secret session key is usually negotiated using the public keys of the two principals.
Virus	A self-replicating code segment; viruses may or may not contain attack programs or trapdoors.
Vulnerability	A weakness that can be exploited to develop an attack against the system or the type of protection that a countermeasure is to provide.
Worm	Propagate itself through memory or networks without necessarily modifying programs. A worm is similar to a virus, because it has the ability to replicate, but differs from a virus in that it does not seek a host.
Wide Area Network (WAN)	A data communications network that spans any distance and is usually provided by a public carrier. Users gain access to the two ends of the circuit and the carrier handles the transmission and other services in between.

# Appendix C

## Characterization of Customer Community Networks

---

- C-1: Public/Commercial Networks (Satellites)
- C-2: Public/Commercial Networks
- C-3: DoD Networks
- C-4: Networking Technologies

**UNCLASSIFIED**

Appendix C  
IATF Release 3.0—September 2000

**Table C-1: Public/Commercial Networks (Satellites)**

<b>Acronym</b>	<b>Full Name</b>	<b>Used By</b>	<b>Purpose</b>	<b>User Bandwidth</b>	<b>Multiple Access Methods</b>
Globalstar	Globalstar Partnership	International long-haul communication. vehicle op. Traveling national and international.	LEO satellite; digital; wireless telecom; extend terrestrial cellular	Digital voice and data; 2400-9600 BPS	CDMA/FDMA
Iridium	Iridium	Cellular telephone market focus	LEO satellite; provide use of portable satellite phones	2.4 KBPS- 4.8 KBPS	TDMA/FDMA
AMSC	American Mobile Satellite Corp.	US, Puerto Rico, V.I., 200 miles of coastal waters; land mobile, fixed site and aeronautical users	Mobile communications for voice, data, digital broadcast dispatch	1.2 KBPS - 4.8 KBPS	FDMA
Ellipso	Ellipsat	Global mobile communication. market	LEO; global voice and data services for access to unserved and remote areas	300-9600 BPS	CDMA
Inmarsat	International Maritime Satellite Org.	Global personal mobile satellite communication.	MEO; satellite communication for commercial emergency and safety app.	For data- 2.4 KBPS - 9.6 KBPS; most terminals 4800 BPS; 64kbps possible in the uplink direction only	TBS
ORBCOMM	Orbital Communications Corp.	Real-time mobile two-way data and messaging services worldwide for US Armed Forces, Transportation, Utility, Oil and Gas	LEO; provides wireless e-mail, Fax, and GPS; small fleet and high value asset location and alarm monitoring use	Iplink 2.4 kbps; downlink 4.8 kbps	TBS
Teledesic	Teledesic	Wireless broadband data app.; also supports voice	LEO; voice, data; high data rate fixed and mobile services; Global coverage	Broadband-155.52 MBPS; capacity for 100k basic 16 KBPS channels	TDMA/FDMA
Odyssey	Odyssey World-wide Services	Global coverage	MEO; wireless personal communication.	4.8 KBPS	CDMA

**Table C-2 Public/Commercial Networks**

<b>Acronym</b>	<b>Full Name</b>	<b>Protocol Authority</b>	<b>Used By</b>	<b>Purpose</b>	<b>Security</b>	<b>User Bandwidth</b>
Internet	Internet	Internet Engineering Steering Group (IESG)	Worldwide	Voice, video and data applications	N/A	300bps - 1.544Mbps
ATM	Asynchronous Transfer Mode	ATM Forum	Data carriers	Wide area networking	data privacy	2.4 GBPS typical
SONET/SDH	Synchronous Optical Network Synchronous Digital Hierarchy	Bellcore/CCITT; ANSI T1.105, T1.107; ANSI T1.106 and ANSI T1.117, ITU-T	US European Japan	Transport of many digital signals w/ different capacities	N/A	51.48 MBPS up to 2405.376 MBPS
PSTN	Public Switched Telephone Network	ITU-14 pending	Worldwide	Provides a wide variety of voice and data services	N/A	9.6Kbps - 155 MBPS



**UNCLASSIFIED**

Appendix C  
IATF Release 3.0—September 2000

**Table C-3 DoD Networks**

<b>Acronym</b>	<b>Full Name</b>	<b>Managed By:</b>	<b>Used By</b>	<b>Purpose</b>	<b>Security</b>	<b>Bandwidth (For User Services)</b>	<b>Nature Of User Access To Network</b>
SIPRNET	Secret IP Router Network	DISA	U.S. DoD	Transport of classified and mission critical data for DOD users	Network operates system high Secret; links encrypted	64 KB - 1.544 MB	Direct network access
ATDNet	Advanced Technology Demonstration Network	DISA	TBS	Research network with various nodes	TBS	TBS	TBS
NIPRNet	N-Level Internet Protocol Router Network	DISA	U.S. DoD	Transport of official data for DoD users	Network operates FOUO. Direct connected to the Internet	Up to T-3 rates	Direct network access and mobile dial-in support
DISN	Defense Information System Network (Superset of SIPRNET and NIPRNET)	DISA	Global DoD community	Primarily data transport system for DII Includes some voice and video	Multiple networks running system high w. cryptographic separation via Trunk Encryptors	Currently as high as DS3 with pressure to increase capacity	Direct and remote dial-in
S-ATM (DAS-C)	DISN ATM Wide Area Network	DISA	DoD community	Secret ATM network; transport of high speed/bandwidth mission critical info.	Encryption	T1 up to OC-3 rates	Via agency network(s)
U-ATM (DAWN; DAS-U)	DISN ATM Wide Area Network	DISA	DoD community		N/A	T1 up to OC-3 rates	Via agency network(s)
JWICS	Joint Worldwide Intelligence Communications System	DIA	DoD community and civilian government agencies	Data and video transport; video teleconferencing at SCI level	link encryption	56 KBPS up to T1 rates	Through agency network or fixed facility
ATMAII	TBS	TBS	TBS	TBS	TBS	TBS	TBS

Acronym	Full Name	Managed By:	Used By	Purpose	Security	Bandwidth (For User Services)	Nature Of User Access To Network
DSINet	Defense Simulation Internet	TBS	TBS	TBS	TBS	TBS	TBS
DREN	Defense Research and Engineering Network	DARPA and DISA	TBS	TBS	TBS	155-622 MBPS	TBS
DRSN	Defense Red Switch Network	DISA	global DoD community	Primarily classified voice system	Trunk encryption	TBS	Phone set
Red IDNX	Red Integrated Digital Network Exchange	USAF	DISN	Data aggregation to achieve bandwidth efficiencies	Trunk encryption	T1 and below	Integrated into transport system
Black IDNX	TBS	TBS	TBS	TBS	TBS	TBS	TBS
MILSTAR	Military Strategic Tactical Relay System	Operated by Air Force Space Command; owned by US Space Command	US military	Emergency action message dissemination; provides tactical survivable communication. including MSE range ext.	Encryption	75 BPS - 2.4 KBPS; medium data rate capability 4.8 KBPS - 1.544 MBPS	MILSTAR terminal

**Table C-4 Networking Technologies**

<b>Acronym</b>	<b>Full Name</b>	<b>Protocol Authority</b>	<b>Used By</b>	<b>Purpose</b>	<b>Security</b>	<b>Bandwidth</b> (For User Services)	<b>Multiple Access Method</b>
ISDN	Integrated Services Digital Network	ITU-T, Q.921, Q.931 ANSI T1.601, T1.408; Bellcore SR 3875	Worldwide	Network browsing; transferring data, remote access	N/A	64000 BPS - 1.544 MBPS	N/A
SONET/SDH	Synchronous Optical Network / Synchronous Digital Hierarchy	Bellcore/CCITT; ANSI T1.105, ANSI T1.106 and ANSI T1.117, ITU-T	US European Japan	TBS	TBS	51.48 MBPS up to 2405.376 MBPS	N/A
ATM	Asynchronous Transfer Mode	ATM Forum	Data carriers	Wide area networking; multi-media and video applications	data privacy	2.4 GBPS typical	N/A
AMPS	Advanced Mobile Phone Service	EIA/TIA-553 US	US	Analog voice	N/A	up to 13 KBPS actual data throughput using CDMA tech.	FDMA
PCS	Personal Communications Service	TIA IS-136A/137A/138A; J-Stds-009/010/011 ETSI PCS 1900, (PCN) DCS 1800	US Europe	Voice and data services; paging-type services; mobile communications	Authentication & Privacy	7.95 KBPS- 13 KBPS	TDMA FDMA
GSM	Global System for Mobile Communication	European Telecommunications Standardization Institute (ETSI)	European	Digital voice, data, SMS	Authentication & Privacy	9600 BPS	TDMA FDMA
DCS	Digital Cellular System	TIA IS-95; IS-136A (D-Amps)	US	Digital voice, data	Authentication	7.9 KBPS	TDMA FDMA
Frame Relay	Frame Relay	Frame Relay Forum (NNI std)-future	US; corporations	data communications; some voice and video	N/A	56 KBPS - 1.544 MBPS	TBS
SMDS	Switched Multi-megabit Data Service	compatible w/ IEEE 802.6 MAN and B-ISDN	Local Exchange Carrier customers requiring large communications pipeline	large data transfers; video, graphics, CAD/CAM, x-rays	N/A	1.544 MBPS - 45 MBPS	TBS

# Appendix D

## System Security Administration

### Duties of the Security System Administrator (SSA)

The SSA must be extremely knowledgeable about the configuration of the system, inherent security weaknesses in the use of the system components, and the security policy. To the extent that a potential threat could exploit the system, the SSAs must also remain current on vulnerability discoveries that may impact their system. The security aspects of the SSA's job are as important to the mission as the operation of the system. To that end, adequate resources must be available to allow SSAs to monitor any security policy violations as well as operational updates.

The SSAs must remain current in relevant technologies (Operating System (OS), audit trails, configuration, known vulnerabilities, etc), and be provided the opportunity to remain current regarding potential attacks on their system. The System Administrator (SA) must keep the system up and running while the SSA ensures the Security Policy is upheld. If there is a security office for the information systems, then a SSA should be a member of that staff.

Under the direction of the Security Policy, the SSAs must operate and at times set up a secure system through use of mechanisms such as passwords (including provisions for protection, distribution, storage, length of character set, and valid duration period of password), security banners that cannot be altered by a user, session controls, lock screen, software and OS patches and updates, and account management. The SSAs must also remain current vis-a-vis potential weaknesses in the system by monitoring appropriate articles and web sites, and they should also be on distribution for OS patches/releases, and Computer Emergency Response Team (CERT) advisories. The SSA's responsibilities include conveying this information to the users, sending advisories, implementing patches, and updating procedures as needed to mitigate risk.

### Configuration Management (CM)

There should be a Configuration Management Plan that includes a CM Control Board (with a security advocate); procedures for access and changes to hardware, software, firmware; detailed and complete system diagrams; a complete map of the system including which ports are available; how the computers in the system communicate with each other; a discussion on who has what privileges; virus protection; Internet downloading and personal software rules; software licensing agreements and procedures; a complete list of system resources (held by the SSA), future requirements; upgrades planned, designed, and proposed; and movement of hardware. The SSAs must remain current on the configuration and is responsible for all upgrades and changes ensuring they do not violate the Security Policy.

# Appendix E

## Office of the Secretary of Defense (OSD) Information Assurance (IA) Policy Robustness Levels

According to the OSD Global Information Grid (GIG) policy, technical IA solutions in the defense-in-depth strategy will be at one of three defined levels of robustness: high, medium, or basic, corresponding to the level of concern assigned to the system. The three levels of technical robustness solutions identified in the OSD GIG Policy are described in the following subparagraphs.

- 1) High robustness security services and mechanisms provide the most stringent protection and rigorous security countermeasures. High robustness solutions require all of the following:
  - National Security Agency (NSA)-certified Type 1 cryptography (algorithms and implementation) for encryption, key exchange, digital signature, and hash
  - NSA Type 1 cryptographically authenticated access control (e.g., digital signature, public key cryptography based, challenge/response identification and authentication)
  - Key Management:
    - For symmetric key, NSA-approved key management (production, control and distribution.)
    - For asymmetric key, Class 5 Public Key Infrastructure (PKI) certificates and hardware security tokens that protect the user's private key and crypto-algorithm implementation.
  - High assurance security design, such as specified by NSA or the International Common Criteria (CC) at a minimum an Evaluated Assurance Level (EAL) greater than 4.
  - Products evaluated and certified by NSA.
- 2) Medium robustness security services and mechanisms provide for additional safeguards above the DoD minimum. Medium robustness solutions require, at a minimum, all of the following:
  - National Institute of Standards and Technology (NIST) Federal Information Processing Standard (FIPS) validated cryptography (algorithms and implementation) for encryption, key exchange, digital signature, and hash (see algorithms at Table 5-4)

# Appendix F

## Executive Summaries

This IATF section is a repository for Executive Summaries. An Executive Summary captures the essence of a user's need in clear, concise statements. The security approach outlined in the Executive Summary points to supporting documentation such as protection profiles.

- *Title - NSA Security Guidance for "Descriptive Name"*
  - *Target Environment*
  - *Potential Attacks*
- *Security Policy and Objectives*
- *Recommended Approach*
  - *Security Functions*
  - *Assurance Requirements*
- *Interoperability Requirements*
- *Supporting Infrastructure*
- *Administrative Information*

**Figure F-1 Executive Summary Outline**

The Target Environment Section describes the purpose and scope of the executive summary and associated protection profiles. It includes the following.

- Which kind of Protection Profile (PP) is this – defense-in-depth, technology goal, customer-specific, etc.?
- Describe the types of user organizations in the scope of this document.
- What does the user organization want the system to do?
  - What is the problem the system is intending to solve?
- Summarize the system environment:
  - Where does the system operate?
  - How will the system be used?
  - Diagram of the system context.

The Potential Attacks Section includes the following.

- What are the information system attacks/events for which protection is needed?  
How can an adversary harm the user organization's mission by attacking the system?  
What non-malicious events (flood, user error, etc.) can harm the user organization's mission through information system effects?
- Attacks should be relevant to the technology under consideration, but should not assume implementation details.

The Security Policy and Objectives Section includes the following.

- What is the organization policy or other rules that the system must meet or support?

Provide the technology-unique context for the policy and objectives (e.g., defend-the-enclave, tunneling).

Referencing GIG (Global Information Grid) policy, describe the robustness category (basic, medium, or high) and any recommended deviations from the policy.

- Describe level of threat and value of information.
- What are the information domains of interest?

An information domain is defined by a *type of information* and the *set of users* with *specific privileges* for access to that information.

- What security objectives must the system meet in order to protect against the information system attacks?

The Recommended Approach Section includes the following.

- What is the conceptual architecture for the system?
- Which security functions are allocated to the technology under consideration?
- What are the dependencies on security functions of other system components?
- Diagram of the system should be included

The Security Functions Section includes the following.

- What are the security functional requirements for the system?  
Include strength of mechanisms and cryptographic algorithm suite.
- What security services must the system perform for each information domain? (e.g., confidentiality, integrity, availability).
- Describe compliance with GIG policy for placement of security functions.

The Assurance Requirements Section includes the following.

- Indicate the required Evaluated Assurance Level (EAL) as defined in the Common Criteria.
- Describe additional assurance requirements or (e.g., Federal Information Processing Standard (FIPS) 140-1 verification).
- Describe compliance with GIG policy for assurance.

Interoperability Requirements Section includes the following.

- What are the interoperability requirements that the system components must meet? (e.g., Transmission Control Protocol(TCP)/Internet Protocol (IP), security protocols)

The Supporting Infrastructure Section includes the following.

- What support does the system require from key management infrastructure (e.g., certificate class and version)?
- What support does the system require from network security management infrastructure (e.g., audit analysis)?

The Administrative Information Section of each Executive Summary must include the following.

- List of Protection Profiles within the scope of the Executive Summary.
- Date and Version Number.
- Author block.
- Approval block.

NSA will provide additional configuration management guidance.



**This page intentionally left blank**

- NIST cryptographically authenticated access control (e.g., digital signature, public key cryptography based, challenge/response identification and authentication)
  - Key Management:
    - For symmetric key, NSA-approved key management (production, control and distribution.)
    - For asymmetric key, Class 4 PKI certificates and hardware security tokens that protect the user's private key
  - Good assurance security design such as specified in CC as EAL3 or greater
  - Solutions evaluated and validated under the Common Criteria Evaluation validation scheme or NSA
- 3) Basic robustness solutions are equivalent to good commercial practice. Basic robustness require, at a minimum, all of the following:
- NIST FIPS validated cryptography (algorithms and implementation) for encryption, key exchange, digital signature, and hash (see algorithms at Table 5-4)
  - Authenticated access control (e.g., digital signature, public key cryptography based, challenge/response identification and authentication or pre-placed keying material)
  - Key Management:
    - For symmetric key, NIST-approved key management (production, control and distribution.)
    - For asymmetric key, Class 3 PKI certificates or pre-place keying material. See reference (p) for policy on migration to Class 4 certificates and software tokens (private keys held in software on the user's workstation.)
  - CC Evaluation Assurance Level (EAL) 1 or greater assurance
  - Solutions evaluated and validated under the National Information Assurance Partnership (NIAP) Common Criteria Evaluation Validation Scheme or NSA

The OSD GIG Policy indicates that the robustness of a network solution must be considered in the context of defense-in-depth and the threat environment in which the system operates. For instance, a system operating on a protected backbone between secure enclaves may not require additional mechanisms for authentication and access control. In addition, if community of interest separation is provided through encryption, it will require less robust solutions.